DOCUMENT RESUME

ED 327 202 IR 053 379

TITLE	Communications in Support of Science and Engineering.
INSTITUTION	Council on Library Resources, Inc., Washington, D.C.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	Aug 90
NOTE	163p.
PUB TYPE	Collected Works - Conference Proceedings (021) -- Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)
EDRS PRICE	MF01/PC07 Plus Postage.
DESCRIPTORS	*Communications; Computer Networks; Correlation; *Engineering Technology; Information Dissemination; *Information Management; Information Technology; Library Collections; Library Services; *Productivity; Publishing Industry; *Researchers; Scientific and Technical Information; *Scientific Research; Tables (Data)
IDENTIFIERS	*Information Infrastructure

ABSTRACT

The papers in this collection, which explore selected aspects of scientific and engineering communication, are the results of a study designed to learn more about the relationships between information resources and scientific productivity. The first three sections contain three key papers related to a conference sponsored by the Council of Library Resources (CLR) that enabled a group of scientists and academic leaders to consider the general topic. A background paper, the transactions of the conference, and a summary statement, taken together, open up many of the pertinent issues and provided a base for future action. Three additional papers, triggered by the conference discussion, were commissioned by CLR: (1) "The Users and Uses of Scientific Information Resources: Recommendations for Study (Helen H. Gee); (2) "Library Resources and Research Productivity in Science and Engineering: Report of a Pilot Study" (Nancy Van House); and (3) "In Preparation: A Comparison of Costs of Library and Information Services in Academic and Industrial Libraries That Support Science and Engineering" (David Penniman and Jay Lucker). Three appendices contain the questionnaire and cover letter for a survey of selected Association of Research Libraries member libraries; statistics on library materials prices; and a table showing the correlations of science library resources with productivity measures by year from 1985 through 1989. (MAB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
E.B. Timmer
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

Communications in Support of Science and Engineering

A Report to the National Science Foundation
from the Council on Library Resources

August 1990

Council on Library Resources
1785 Massachusetts Avenue, N.W., Suite 313
Washington, D.C. 20036

Preface

During the summer of 1989, the National Science Foundation asked the Council on Library Resources to explore selected aspects of scientific and engineering communication with the objective of learning more about the relationships between information resources and scientific productivity. Such information is of use in the preparation of the Science and Engineering Indicators series of the National Science Board. The enclosed papers, which constitute the report of the project, are the result of CLR's effort. The topic is a complex one, but it is the Council's hope that what has been done will point the way toward productive areas for investigation.

The key papers are related to a conference sponsored by CLR that enabled a group of distinguished scientists and academic leaders to consider the general topic. A background paper, the transactions of the conference, and a summary statement, taken together, open up many of the pertinent issues and provide a base for future action. The recommendations in the transactions are especially important.

Three additional papers, triggered by the conference discussion, were commissioned by CLR. The first, by Helen Gee, considers how more might be learned about the information needs and information-seeking methods of scientists and engineers. The second, by Nancy Van House, explores, in a very preliminary way, whether a relationship exists between the extent of library resources and services and the quality and distinction of scientific research in academic settings. Finally, the third study, now in preparation at Bell Laboratories and the Massachusetts Institute of Technology, will compare the investment made for information resources in a small set of academic science libraries with that made for comparable purposes in a set of industrial libraries. The study will be distributed in the fall of 1990.

During the past three or four years, CLR has made a number of grants and sponsored many meetings, all with the purpose of learning more about information needs in all fields and promoting among faculty, librarians, and university officers the kind of open discussion that is essential to long-range planning for information services. While the concerns of scholars working in humanistic and historical disciplines differ from those of scientists, there are many similarities in their information settings. For both groups, the same forces are at work--ever-increasing costs, the rapid development of information technologies, and the sheer quantity of information that must be absorbed into the system. In the context of library operations, the principal differences between science and the humanities are the dependence on current, versus historical, information and the requirement for speed of access. Even so, some areas of science need full access to both the historical record and the most current information. The sciences, by their nature, need to have their information base analyzed in great detail, but, increasingly, humanists are looking for comparable attention to their literature. Further complicating matters is the requirement to meet various needs for scientific and technical information--for research, obviously, but also for many commercial and industrial enterprises, for teaching at all levels, and for the general public, broadly defined.

All this is to say that the search for signs of scientific progress in the information structure is difficult. It is a given that future scientific progress is based on information and the product of science is new information. It seems reasonable to assume that the effectiveness of the information "system" affects the performance of science itself. In turn, shaping the ideal system and measuring its effectiveness in absolute terms is probably an impossible assignment, especially at this point in time when definitions of "ideal" are largely personal and all system elements are in a state of flux.

But even given the difficulty of understanding the relationship, its study seems worthwhile and carries the promise of improvement that comes with understanding. From what we have learned during the past year, there are three basic aspects of this general inquiry that merit much more thoughtful and imaginative attention. By their nature, they require more than routine analysis.

1. The future form of scientific publishing

   How will the growing trend to integrate the information structure with the research process affect the traditional system of scientific publication? Publication in established journals still provides the authenticated historical record for science. However, escalating costs, the uncontrolled national and international commercialization of scientific publishing, and the sheer quantity of material offered for inclusion put the system--which is based on the assumption of wide distribution--seriously at risk. Given the capabilities of information technology, the realities of funding, the need to protect the authenticity of scientific information, and the requirement for unconstrained access, what should be the future form of scientific publishing?

2. The characteristics of and requirements for scientific communication

   There are various forms of scientific communication, some essentially personal, others affected in various degrees by external conditions. Because of computer, telecommunications, and information storage technologies, the options for communication among researchers are both greatly expanded and more complex. Too little is known about the needs of scientists and engineers for information, about methods of information transfer related to research needs, and about variations in ways of working and what determines those variations. Further, a better understanding of the relationship between the form and characteristics of information and the utility of such information is essential to the design of future information services. Special care not to assume homogeneity of need is essential in any investigation.

3. The future form of library services and information systems for science

Research libraries are being redefined. Rather than being simply buildings, collections, and staff, they are becoming organizations that are intricately linked to each other for purposes of identifying, locating, and providing information in all forms--publications, images, sound, and data. They have made great strides in applying information technology to operations, they are increasingly interdependent in building and maintaining collections, and they have staffs that are more diverse than ever, with technological skills, subject knowledge, and management abilities all represented.

However, the process of making fundamental change, given operating realities and still valid traditional responsibilities, is difficult. Success will depend in large part on constructive collaboration with user communities. A purposeful and long-term effort to recast library services for the sciences is required. Among matters requiring exploration is the need for a National Library for Science, as one element in a system to assure a comprehensive information service for the sciences. To this end, the function, structure, and organization of a new national library or service (perhaps one that brings together in some effective combination the Library of Congress and other key science libraries) should be described, paying special attention to the relationship with academic research libraries, prospects for institutional cost containment, and service enhancement for science.

These three topics seem to us to be the key matters for further attention if we are not only to understand how science information and science itself interact, but to make certain that the interaction is as efficient and as productive as possible.

Finally, we want to record the contribution of Dr. Martin M. Cummings, Consultant to CLR for managing this project; the members of the CLR staff who assumed various responsibilities; and the many individuals who took part in the discussions that helped shape this report.

Warren J. Haas

August 1990

Communications in Support of Science and Engineering

CONTENTS

Section I.

Summary (Martin M. Cummings)

Section II.

Discussion Paper for Conference (Martin M. Cummings)

- A. Background
- B. Influence of New Information Technologies on Research
- C. Elements of the Existing Information Infrastructure
- D. Considerations for Improving the Information Infrastructure
- E. Some International Implications
- F. A Proposal to Improve Library Services to Science and Engineering

Section III.

Transactions of the Conference on Communications in Support of Science and Engineering (Eleanor W. Sacks)

Section IV.

Special Studies Resulting from the Conference

- A. The Users and Uses of Scientific Information Resources: Recommendations for Study (Helen H. Gee)
- B. Library Resources and Research Productivity in Science and Engineering: Report of a Pilot Study (Nancy Van House)
- C. In preparation: A comparison of costs of library and information services in academic and industrial libraries that support science and engineering (David Penniman and Jay Lucker)

Communications in Support of Science and Engineering

Section I

SUMMARY

Martin M. Cummings, M.D.

In a society that is increasingly influenced by advances in science and technology, it is imperative that we understand the processes through which scientists and engineers produce, acquire, and transmit scientific information. In response to a request by the National Science Foundation, the Council on Library Resources sponsored a conference designed to identify and analyze the major trends in the use, storage, and dissemination of scientific and engineering knowledge and to provide an analytical framework for possible use in future issues of Science and Engineering Indicators, the biennial NSF publication that uses quantitative measures to reveal trends in science and engineering in the United States. This meeting brought together senior university officers, information and communications specialists, librarians, and editors as well as scientists and engineers. Based on presentations and discussions, a number of important trends and problems were identified, which are summarized below.

New technologies increasingly shape the way science and engineering are performed. Computers, in particular, have made it possible to attack scientific problems that could not have been pursued without their ability to create, store, and manipulate data. At the same time, computer-based communications networks provide rapid and efficient mechanisms for information transfer. However, as yet there has been no significant effort to better understand quantitatively how the information resources of the nation serve its research enterprise. These resources are located in the nation's universities, industries, and its federal government. They consist of information conveyed through publishing, meetings and conferences, and electronic networks as well as many other informal means of communication. The underlying reason for seeking a better understanding of the information infrastructure in the United States is the commitment to applying new technologies to improve research, education, and service in our nation.

Publications and New Alternatives

Formal publication remains the most common form of scientific reporting, providing a large part of the information base for new research and for the dissemination of research findings and results. Academia and granting agencies have used the number and quality of publications as a measure of a researcher's creativity and productivity. Citation indexing of published articles is used by NSF as an indicator of the importance of scientific publications by discipline or field of interest.

The volume of the scientific literature continues to double in size approximately every twelve years, threatening to overwhelm our ability to store, search, and retrieve information. The publication process, which increasingly is managed by commercial publishers, is under stress. The rapid growth of publications has been accompanied by an alarming rise in the cost of scientific serials, which has already constrained the ability of U.S. libraries to acquire and maintain important collections. It is alleged that these price increases result in large part from the takeover of U.S. publishers by large foreign conglomerates. The volume of publication, its high cost, and the adverse effects on libraries that index, catalog, bind, and store printed materials have led to a search for improved methods of processing scientific information.

New technologies are changing the current publishing process. Computer-generated desktop publishing allows authors to create text and graphics that can be transmitted easily among scientists and engineers. This capability has the potential to reduce dependence on the existing publications system. The movement toward electronic publication is driven largely by economic considerations, but it also reflects a desire to speed up the transmission process. However, it is still unclear whether this new mode of communication will be susceptible to the bibliographic control necessary to organize, find, and retrieve articles published by this method. Desktop publishing also lacks an established mechanism for refereeing and other quality controls that are in place in the printed journal literature.

Libraries

Research libraries acquire, organize, and maintain the collections of record for all fields of science and engineering, including publications in all formats. Increasingly, libraries maintain scientific databases that are available to users upon demand. Although it is difficult to quantify the role of the library in the information infrastructure supporting science and engineering, it has been proportionately a large component. The 119 members of the Association of Research Libraries spend approximately $1.5 billion annually. Between 1985 and 1988, the expenditures for library materials increased by 30 percent, while the number of volumes declined by 10 percent. Approximately 10 percent of library budgets are utilized for automation of records and services. This expenditure is expected to increase with time.

It has proved difficult to correlate the size of library collections with the productivity of research. However, in a general way it appears that leading research institutions maintain strong research libraries to support their educational and research activities. These libraries have created consortia to make available, through cooperative cataloging and interlibrary loan, materials that may not be easily found in an individual library. They also utilize computer systems for bibliography and reference, providing information efficiently on demand.

Although research libraries have made progress in applying new information technologies to improve services, the rapid growth of science and engineering has made it difficult for libraries to maintain comprehensive support for research and development in all of the sciences. New organizational and technical arrangements need to be explored to provide enhanced services for science and engineering.

Biomedicine and agriculture have national libraries that serve many of the information needs of their scientists and researchers. A planning effort involving scientists and representatives of science libraries should be encouraged to study how comparable services can be provided to all disciplines of science and engineering. For example, scientists and engineers would benefit greatly from the creation of a National Library for Science and Engineering. Such a library might be based on the model of a central library, such as the Canada Institute for Scientific and Technical Information, or it might be created through a consortium of existing strong science libraries. This organization could act as a National Periodicals Center, which would serve as a comprehensive source of published materials that could be shared with academic science centers, research institutes, professional associations, individual scientists and engineers, and industry. The availability of such a national resource would alleviate the problem of declining science literature collections in our nation's libraries.

With such a central facility available, procuring information as a scientist needs it rather than buying large collections of literature in anticipation of need may become the most efficient and cost-effective strategy for university libraries striving to meet information needs. This benefit has been made possible by the development of electronic networks and the telefacsimile transmission of documents.

Computers and Networks

Computer-based networks play an increasing role in recording and transmitting research data. In some fields, the amount of data accumulated exceeds the capacity of existing information storage and retrieval systems. It has been estimated that five hundred databases, or 30 percent of all databases, are related to science, engineering, and medicine. More than six hundred networks with over 100,000 computers are linked through a consortium of networks called INTERNET. Rapid growth in the use of this network, now approaching saturation, suggests that a larger and faster network will be needed in the near future. The recent proposal in Congress to create a National Research and Education Network is designed to greatly increase the efficiency and capacity through which information can be accessed and transmitted.

The rapid extension of the use of electronic mail has had a major impact on the way scientists and engineers communicate. These changes in information seeking are being further accelerated by the rapid introduction of telefacsimile as a means for transmitting documents. Other transmission media, including optical disks and CD-ROM, increasingly include scientific information for individual access and use. In addition, many highly specialized databases in narrow fields of science can be shared by transmission of tapes or disks to small and large research departments or institutes.

During the past decade, there has been a trend toward direct user access to online databases. End-user searching has been stimulated by improved software that makes access more user friendly. The costs of searching and document delivery are a major factor in determining use. Universities may need to subsidize access to important databases by graduate students, faculty, and other individuals who do not have funds for this purpose.

Suggested Areas for Further Study

The CLR/NSF conference clearly indicates a need for a well-designed study of the current information-seeking behavior of scientists and engineers. This effort is deemed an essential element for strategic and technical planning in the development of more efficient information systems. It is also believed that such a study will contribute to the search for new indicators to measure the role of libraries and other information systems serving science and engineering.

It is suggested that NSF consider the feasibility of creating a National Library for Science and Engineering utilizing the existing resources within the nation's leading research libraries. This will require a study of organizational arrangements and economic and financial considerations, as well as an analysis of the information resources needed to provide prompt, comprehensive, and reliable services to the science and engineering communities.

Based largely on conclusions reached at this conference, CLR commissioned several exploratory studies (see Section IV) that may help NSF identify new indicators of productivity in U.S. science and engineering by focusing attention on the relationship between measures of library performance and outputs of science. It is important to understand better the contribution of the research library to the research enterprise. Also, it is important to learn more about the characteristics of scientists' use of libraries, particularly the number and kinds of uses made of library services through online networks.

One such study, by Nancy Van House, University of California, Berkeley, explores the relationship between information resources and research accomplishments. Specifically, data are provided that allow the measurement of library resources and services utilized by academic scientists. In a separate but related study, David Penniman, Director, Libraries and Information Systems, AT&T Bell Laboratories, and Jay Lucker, Director of Libraries, Massachusetts Institute of Technology, compare the costs of library and information services in support of science and engineering performed in academic and industrial libraries. The Council also commissioned a paper that provides a comprehensive approach to the design of a national study describing the information-seeking behavior of scientists and engineers. A group of communications and information specialists met in February 1990 to assist in this effort. Helen H. Gee, former Chief of the Office of Program Evaluation at the National Institutes of Health, is the author of this paper.

We believe the papers that constitute this report will be useful to leaders of science and engineering in planning communications and information systems for the future. They do not cover all of the issues that need attention, but they serve as a backdrop for future work designed to improve scientific information transfer. The observations of the participants reflect the need for further analytic studies to provide information to the people in federal agencies and private industry who are concerned with the creation of a national science information system. A better understanding of the relationship between the forms, characteristics, and uses made of scientific information is needed for the design of such a system.

Communications in Support of Science and Engineering

Section II

DISCUSSION PAPER FOR CONFERENCE

Martin M. Cummings, M.D.

A. Background

The National Science Foundation (NSF) asked the Council on Library Resources to organize a process through which existing and additional indicators might be used to identify major trends in the organization, storage, dissemination, and use of scientific and engineering information. NSF also asked that the process serve as a base for analytical studies in future publications of Science and Engineering Indicators. This objective requires an understanding of how the existing information infrastructure contributes to the quantity and quality of science and engineering in the United States.

Leading scientists, engineers, librarians, and information scientists were invited to a conference to describe current information sources, the information-seeking patterns of scientists, and changing communications and information systems that significantly affect the performance of science and engineering. Issues and problems requiring further study were identified as a basis for commissioning special reports by experts selected for this purpose. A list of conference participants is found at the end of Section III.

Since 1972, the National Science Foundation has published a biennial series of reports that describe and measure changes in the national enterprise that encompass scientific research, technology, and engineering. Originally entitled Science Indicators, these reports have more recently been issued as Science and Engineering Indicators.[1]

In accordance with the National Science Foundation Act of 1950, these reports are transmitted to the President and the Congress to provide a comprehensive set of quantitative data that may be valuable in studying trends in science and engineering and in developing public science policy. The National Science Board provides oversight and guidance to the Division of Science Resources Studies, whose staff carries out the data collection and analysis that leads to the preparation and ultimate publication of the report.

Science and Engineering Indicators provides an impressive amount of quantitative data about the organization and performance of U.S. science and technology. An effort is made to provide this information in the context of a changing social, educational, and research environment. Much of the information relates to the following elements of the U.S. science and engineering system as characterized by NSF.

---

[1] National Science Board, Science and Engineering Indicators (Washington, D.C.: U.S. Government Printing Office, 1987).

1. Human resources available for research, technical support, and management.
2. Organizational arrangements for research and development.
3. Physical infrastructure for research and education.
4. Financial support available for the research enterprise.
5. Major ideas, research methods, and strategies embodied in the science and engineering literature.

B. Influence of New Information Technologies on Research

In a paper entitled Science and Technology in the Academic Enterprise: Status, Trends, and Issues, the Government-University-Industry Research Roundtable (sponsored by the National Academy of Sciences and the National Academy of Engineers) described the changing and increasingly complex research environment.[2] The report indicated that new technologies increasingly shape the scholarly agenda in the sciences and engineering. New types of instruments provide previously unattainable precision and scale of experimentation. Computers make possible large-scale data analysis, while computer-based communications networks provide efficient and rapid mechanisms for information transfer.

While it is recognized that these changes have affected the performance and productivity of science, it has been difficult to isolate or identify an indicator that measures these effects qualitatively. The number of publications and patents produced by researchers appears to be the most reliable indicator of the activity of the research enterprise. Such indicators may not, however, reflect its quality, although some believe that citation indexing and analysis serve this purpose. In some fields, the productivity per scientist appears to have increased after more sophisticated instrumentation and electronic information networks became available. Although these developments differ by field, it has been suggested that together they have led to more reported research. One recent study indicates that the number of physiology journals is growing at a rate more than double the increase in the number of specialists in this area.[3]

---

[2] Science and Technology in the Academic Enterprise: Status, Trends, and Issues. A discussion paper. The Government-University-Industry Research Roundtable. Washington, D.C.: National Academy Press, October 1989.

The importance of new information technology in generating, analyzing, and transmitting the results of research was recognized by the Institute of Medicine, the National Academy of Engineering, and the National Academy of Sciences when they created a joint Committee on Science, Engineering, and Public Policy to study how new information technology is influencing the conduct of research. In the preface to the report of the Panel on Information Technology and the Conduct of Research, Donald N. Langenberg writes:

---

[3] Eugene M. Renkin, "Increased Cost of Professional Journals: A Faculty Member's Perspective." UC-Davis Library Perspective 6 (May 1988): 5.

If asked to distill one key insight from my service on this panel, I would respond with the assertion that information technology is of truly enormous importance to the research community, and hence to all humanity, precisely because it has the potential to enhance communication of information and knowledge within that community by orders of magnitude. We can now only dimly perceive what the consequences of that fact may be. That there is a revolution occurring in the creation and dissemination of information, knowledge, and, ultimately, understanding is clear to me. It is also clear to me that it is critically important to maintain our commitment to free and unfettered communication as we explore the uses of information technology in the conduct of research.[4]

C. Elements of the Existing Information Infrastructure

As yet, there has been no significant effort to relate all of the information resources of the nation to its research effort. The information infrastructure that supports science and engineering involves the resources of the nation's

---

[4] Information Technology and the Conduct of Research, Report of the Panel on Information Technology and the Conduct of Research, Committee on Science, Engineering, and Public Policy--a joint committee of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine (Washington, D.C.: National Academy Press, 1989), ix.

(1) universities, (2) industries, (3) computer and communications systems, and (4) electronic networks, as well as (5) institutions and mechanisms such as research libraries, the publishing industry, meetings and conferences, and informal means of communication. Unfortunately, many of these activities are poorly linked to others and thus present a less efficient system than might otherwise be available.

Nearly everyone agrees that communication influences the conduct and efficiency of the research process. Publication documents the performance of research, and analysis of the aggregated published output of science can serve as an indicator of trends and achievements in major fields of science. The diffusion of knowledge occurs in many ways, but clearly the publication of research results is a powerful mechanism in this process.

Publications

Publication exists at both ends of the research enterprise: it provides part of the information base for new research and reports research findings and results. Academia and granting agencies have used the number and, less commonly, the quality of publications as a measure of a researcher's creativity and productivity. The growth of the scientific and technical literature has been extremely rapid, doubling in volume approximately every twelve years. It is possible to estimate the size and characteristics of a given field of interest when these publications are aggregated by discipline. Scientific and technical literature may be analyzed by language or country of origin.

In recent years, NSF has used the number and share of articles published by U.S. scientists and engineers in the world's journals as an indication of the relative strength of national science and technology. The number of citations to an article in subsequent publications (i.e., citation index) is used as an indicator of the relative significance of the work. Recently, the U.S. share of publications in leading science journals has fallen as science and technology activity has increased in other countries.[5]

Citation indexing may be viewed as a rough measure of the quality of publications; examined on a recurring basis, citation indexing may be considered an indicator of the merit of contributions to various scientific fields. Analysis of frequently cited publications may serve as a means of detecting changes in direction or emphasis in various fields of research.

In addition, the number of articles published in a given field reflects the relative size of the endeavor. For example, in tabulations reported in the 1987 Science and Engineering Indicators, it may be seen that American medicine (clinical and biomedicine) leads all other fields in the number of publications in international scientific literature.[6] In earth sciences, space sciences, and engineering, the U.S. contributes about 40 percent to the world's literature in each of these fields. To portray accurately the U.S. contribution, it may now be necessary to estimate the output of science that is transmitted or reported through new modes of publishing and communication.

---

[5] National Science Board, Science Indicators: The 1985 Report (Washington, D.C.: U.S. Government Printing Office, 1985).

Electronic Publishing

New technologies are changing the way scientists communicate. The emergence of "desktop" or electronic publishing using new computer technologies may affect the current publishing process. Presently, desktop publishing is largely used to produce preprints. It can be used to produce formal publications, however, given appropriate editing, marketing, and distribution systems. Although electronic publishing emerged as a reality less than a decade ago, it has been estimated that by 1990 it will represent a $6.5 billion market.[7] This may be an important indicator of a changing trend in scientific publication. If this estimate is generally correct, desktop publishing will have a profound impact on the bibliographic systems that attempt to organize the literature for easy access, as well as on libraries accustomed to acquiring literature in conventional published formats.

---

[6] Appendix tables 5-26, 5-27, 5-28, Science and Engineering Indicators, 1987.

[7] Connie Winkler, "Desktop Publishing," Datamation 32 (1 December 1986): 92-94, 96.

The availability of inexpensive computers has provided unprecedented independent computing and publishing capabilities. The quality of the traditional review process may be affected by changes in workflow for editorial services. Text and graphics that can be generated and manipulated easily by scientists and engineers will result in less work for the existing centralized publishing apparatus. However, decentralization runs the risk of creating large numbers of incompatible systems that are difficult to integrate. If scientists and engineers use electronic publishing methods more extensively, the existing publication system will require modification. Thus, it may become important to find a way to measure the influence of these new transactions on the output and productivity of science. The decade-long movement toward electronic publishing is largely driven by economic considerations. Sales of desktop publishing equipment increased from $993 million in 1988 to $1.36 billion in 1989. It is estimated that these sales will reach $5.33 billion in 1993.

A critical review of the problems of control and the potential effects of desktop publishing on faculty use and distribution of information at the University of California, Los Angeles, has been presented by Hayes.[8] Despite these problems, desktop publishing appears to be gaining momentum in many academic and industrial institutions.

---

[8] Robert M. Hayes, "Desktop Publishing: Problems of Control," Scholarly Publishing 21, no. 2 (January 1990): 117-23.

The Economics of Publications

Scientists and engineers, faced by a constantly increasing amount of new information appearing in many formats, have begun to take economic factors into account when evaluating its utility. The value of information is measured in part by its use in minimizing work expended in solving problems. Huth has developed an equation ($\text{Value} = \frac{\text{Utility}}{\text{Cost}}$) to calculate the value of information. In his equation, the value of the numerator, Utility, is determined by the relevance of the information to the user's needs. The denominator, Cost, is determined not only by the purchase or service cost but also by the time needed for access or retrieval.[9] Thus the cost of information may be high if a significant amount of time is required to find it. Also, the cost of information varies depending on the medium used. Huth provides an interesting example of the varying costs of information in printed journals compared to other sources.

Approximate Cost (Sale Price) per 1000 Words for Three Types of Information Sources (1980 Data)

New England Journal of Medicine	$ 0.01
Annals of Internal Medicine	$ 0.02
American Review of Respiratory Diseases	$ 0.05
Audio cassettes (Audio Digest)	$ 0.50
Postgraduate courses	$10.00

---

[9] Edward J. Huth, "The Information Explosion," Bulletin of the New York Academy of Medicine 65, no. 6 (July-August 1989): 653-55.

Scientists make heavy use of journals as a source of information. Libraries, in an effort to serve science, invest heavily in the acquisition of scientific literature. In recent years, the cost of journals has risen rapidly, forcing libraries to drop subscriptions.

In a review of the rising cost of serials, Palmer makes several important observations.[10] First, while prices of European publications are rising most rapidly, American publishers also have raised their prices beyond the actual rate of inflation. For example, a survey of the core journals acquired by the University of California in fifteen disciplines revealed a price increase of 31.9 percent over a two-year period. Second, in reviewing the one hundred titles most frequently cited in Science Citation Index, Palmer found that twenty-eight were commercial publications, while the remaining seventy-two were official organs of professional societies, research associations, or academies. He found no correlation between a journal's perceived research value and its price. The journals published for profit were approximately three times more costly than those published by nonprofit groups. He cites as an extreme example the commercial publication International Journal of Neuroscience, which costs about $1.00 per page, in contrast to the Journal of the American Chemical Society, which costs $0.04 per page.

---

[10] Raymond A. Palmer, "Suggestions for a Partnership," CBE Views 12, no. 1 (1989): 9.

Some heretofore overlooked or neglected problems now emerge as significant issues for examination. Some related examples are:

1. A serious economic problem is thought to result from the takeover of U.S. publishers by large foreign conglomerates. In the sciences, this trend has been shown to result in very large increases in the prices of scientific serials.
2. Many professional and scientific societies have transferred the production of their primary publications to commercial publishers, both domestic and foreign.
3. This trend has led to a situation in which the U.S. taxpayers and industries who supported U.S. research must now pay high prices to obtain the results of this research published in foreign-owned journals. For example, a large university library reported that 3.7 percent of the library's subscriptions come from three commercial publishers and account for 22 percent of its serials budget.

The increasing influence and power of foreign-owned commercial publishers may lead to a situation in which American research libraries are no longer able to acquire and maintain collections that are sufficient to satisfy the needs of faculty and graduate students, the largest group of users of research libraries. This limitation, added to the expansion and fragmentation of many scientific fields (which leads to the creation of new specialty journals), will make it necessary for the research library to face the difficult task of selecting materials that best meet the needs of its constituent user communities. Scientists and engineers need to be called upon to assist in this important selection process.

Patents

Science and Engineering Indicators has used the number of patents taken out on new inventions as an indicator of invention. However, since industries vary considerably in their propensity to patent inventions rather than hold them as trade secrets, comparisons among different industries are not thought to be reliable. NSF has also examined the number of patents issued to foreigners as contrasted with Americans.

Narin and Frame recently described several measures used to delineate the rapid growth of Japanese science and technology during the past decade.[11] They used the number of U.S. patents held by Japanese inventors as an indicator of technological size; the number of patents that are cited by other patents as an indicator of technological impact; the extent to which these patents cite the nonpatent literature as a measure of the linkage between science and technology; the number of papers published in the scientific literature as a measure of scientific size; and the frequency of citation to this literature as a measure of scientific impact.

---

[11] Francis Narin and J. Davidson Frame, "The Growth of Japanese Science and Technology," Science 245 (August 11, 1989): 600-605.

This type of analysis of patents as indicators of scientific and technological strength may be one of the most important measurements to be monitored on a recurring basis. The share of patents held by individual countries may be analyzed by technical classifications to determine growth or regression in a dynamic way.

Preservation of Scientific and Engineering Literature

Scientific publications not only serve to report the results of research, but also may be used as a retrospective tool to trace the origin of ideas and concepts. In the aggregate, they serve as mankind's memory of events that have transpired over long periods of time. Unfortunately, much of the printed record of science is disappearing due to the deterioration of acid paper. As a response to this crisis, the Commission on Preservation and Access was created to coordinate the preservation of the published record in all fields.

The need for historical documents by scholars and scientists varies greatly by discipline. Unfortunately, it is too costly to preserve all deteriorating materials, including machine-readable records as well as conventional printed records. There are significant differences in the priorities for preservation of older scientific manuscripts and publications that are of interest to historians and archivists of science. It is important for leaders of science and engineering to engage in the process of critically selecting materials to be preserved. The selection of the most important resources for future historical studies should involve all public and private organizations engaged in scientific research and development.

Libraries

In the past, some effort has been made to evaluate the strength of the nation's research libraries as an essential resource for research and engineering. The data used for this effort were drawn largely from the annual reports of the Association of Research Libraries (ARL). While ARL data have the virtue of continuity and stability, they fail to provide any quantification of the entire information infrastructure (public and private) that is available to assist American research and educational programs. In large part, the data are not adequately detailed to allow assessment of library efforts in support of even the major subject fields.

It is difficult to assess the role of the library in the information infrastructure. However, it has been a proportionately large component. Today the 119 members of the Association of Research Libraries expend approximately $1.5 billion annually. A recent study of 75 university libraries revealed that between 1970 and 1980 expenditures for library materials increased by 91 percent, while the number of volumes added per year decreased by 23 percent. Between 1985 and 1988, expenditures increased by 30 percent, while the number of volumes purchased declined by 10 percent.

The ARL statistics serve as indices of the size of collections, expenditures, staffs, and selected services that are provided to users. ARL also reports enrollment and Ph.D. data for member universities. These data have been collected (with some changing definitions) for seventy years. Thus they reflect a quantitative measure of the evolving changes in research library efforts to provide services to their users.

It has proven difficult to correlate the size of library collections with the quantity or quality of ongoing research in a given organization. In a general way, leading research institutions maintain good research libraries to support and assist the local scientific enterprise. However, the emergence of library consortia and the expansion of networks that provide ready access to databases and specialized information services located at distant sites provide new mechanisms through which scientists and engineers keep abreast of ongoing research and development.

Further measures of library performance are needed to reflect the contribution of the research library to the research community. These measures might include: access to and relevance of the collection to the science profile of the user community, rather than collection size, as an indicator of utility; amount of money spent for automation and systems development as an indicator of a commitment to improve services;

and efforts made to obtain requested materials through interlibrary loan, as an indicator of the quality of service provided.

It has been estimated that total expenditures on information resources account for 11 to 12 percent of the university budget. Generally, the research library is allotted between 3.5 and 4.5 percent of the total general and educational budget of the parent institution. It is important to examine institutions that are largely devoted to training and research in science and engineering to seek correlations between the strength of the library and the perceived quality of the research programs. Also, it would be interesting to know how scientists and engineers use libraries in their search for information. It would also be useful if libraries and other institutional information providers would report on the number and kinds of uses made of their services. In particular, it would be interesting to know the frequency of use of information in printed form, as well as information available from online computer systems, databases, and such devices as CD-ROM and optical disks. Following trends in the characteristics of information use in different formats should be helpful to science managers in planning and budgeting for information resources.

The National Libraries in Support of Science and Engineering

In the aggregate, the three national libraries represent a major information resource for science and engineering research and development. The National Library of Medicine and the National Agricultural Library provide comprehensive services to their constituencies. The Library of Congress contains a vast collection of publications in many fields of science, but it has not developed a mechanism that would allow these materials to be shared easily with the nation's scientific communities. There is a need for closer linkages between the Library of Congress and academic and scientific organizations. The Library of Congress has recently indicated its intention to explore the need for providing an information service to scientists and engineers. If this development becomes a reality, an important gap in literature services to scientists and engineers could be filled.

The National Technical Information Service is an important special source of scientific and technical information. It maintains a massive collection of technical reports that cover all fields of science and engineering. Its services are provided on a full cost-recovery basis that is unique in government. The material available within the collection comes largely from contractors working for federal agencies, and it also contains a significant number of documents obtained from foreign sources. The collection is used principally by commercial or industrial organizations but should be of value to academic research institutions as well.

It has been suggested that consideration be given to establishing a national organization to provide integrated information services to areas of science not served by the national libraries. The Canada Institute for Scientific and Technical Information is a model for such a development. Another approach would utilize an organizational arrangement through which a consortium of major U.S. research libraries would serve this purpose.

Computers and Networks

Computer networks play an increasing role in research. Very large databases produced by the American Chemical Society, the National Library of Medicine, and many other professional associations are now available to users online. They reflect the massive amount of research output over time. New technologies to record, store, and retrieve information appear with some regularity. In some fields (e.g., satellite and space studies), the amount of data accumulated far exceeds the capacity of current systems for storage and retrieval. In a recent editorial, Philip Abelson described the enormous increase in scientific data resulting from the use of electronic equipment.[12] He estimated that 30 percent of all databases are related to science, engineering, and medicine. The use of computer-based information systems and services allows individuals to access this vast store of information. Currently, more than 600 networks with over 100,000 computers and workstations are linked through INTERNET. Rapid growth in the use of this network suggests that a larger and faster network will be required in the near future. Although it is a difficult task to accomplish, monitoring the flow of computer searches and electronic transmission of data may provide a useful indicator of the volume of research activity in selected fields.

---

[12] Philip H. Abelson, "Retrieval of Scientific and Technical Data," Science 245, no. 4913 (July 7, 1989): 9.

Databases that serve science and engineering continue to grow at a rapid pace. More than five hundred databases are available online through public or private organizations. Bibliographic databases of interest to scientists can now be accessed from many university, industrial, and public libraries. In addition, many highly specialized databases in important but narrow fields of science are maintained by large research institutes or departments.

During the past decade, there has been a trend toward direct patron access to online databases. End-user searching has been fostered by improved software that makes access more user friendly. The increased availability of microcomputers and networks has led to researchers who generate information also becoming direct information consumers without the benefit of intermediaries.

Studies sponsored by the Council on Library Resources reveal that scholars and scientists prefer subject searching. Thus, libraries and database generators have given a high priority to improving and enhancing various subject search systems. Efforts to apply expert systems methodologies to online catalogs offer hope that user-system interfaces will be based largely on natural-language query and response. Because of improved systems for document location and delivery based on facsimile transmission, a scientist can expect to have information rapidly available in the laboratory, office, or home.

The costs of searching and document delivery will be a major factor in determining use. Databases that are generated and made available by public agencies generally are provided at a reasonable fee, while those that are provided by commercial vendors vary greatly in their pricing structures. Society may need to find some mechanism to subsidize access to important databases by students and faculty who do not have funds for this purpose.

D. Considerations for Improving the Information Infrastructure

During the past twenty years, many independent information systems have been created to meet the special needs of various fields of science and engineering. To be effective, the scientist should be able to seek and find information from a single entry point. Standardization of hardware, software, and network development is necessary to meet the goal of single entry point searches. The creation of national communication networks dedicated to research will do much to overcome the incompatibilities that exist among local networks. Electronic transfer of data and messages may now allow information exchange and collaborative studies to take place in an interactive and more productive way. Networks may also stimulate the growth of interdisciplinary research.

Recently, the National Academy of Sciences, EDUCOM, and other interested organizations have recommended that a National Research and Education Network be designed and implemented. The existing NSFNET would become a major element of such a development. NSFNET is expected to be saturated by 1990 and thus will require a major expansion of its capacity. Such a network would allow for interactive communication as well as transmitting massive amounts of data upon request. It is expected that research productivity will be improved by facilitating access to information and collaboration among researchers.[13] To be effective, systems and network designers should take into account the projected information needs of the ultimate users of the network.

The Committee on Science, Engineering, and Public Policy recommended that a users' group be formed to advise those concerned with systems development on the use and evaluation of information technology. This group can be most helpful in the design of a national infrastructure for the use of information technology in research. From such a group, the agencies and institutions that may finance this undertaking will expect reliable estimates of costs, an assessment of possible legal constraints, and the development of hardware and software standards that will make it possible for communication to take place through efficient interconnections. The respective roles of university libraries and computer centers as information service providers will also need to be clarified. At the same time, some accommodation with commercial vendors of information will be required.

---

[13] Douglas E. Van Houweling, "The National Network: A National Interest," EDUCOM Review 24, no. 2 (Summer 1989): 14-18.

The information requirements of scientists and engineers should dominate the final design of the infrastructure to be developed. The future of some forms of research and engineering will depend on the success of such a national network.

E. Some International Implications

As science and engineering have developed in other countries, an increasing amount of information has been published in foreign languages. Translation of such material may be required to make the knowledge generated elsewhere available to U.S. scientists. Large-scale translation programs previously funded by U.S.-owned foreign currencies are no longer available.

There are some government officials who believe that the cutting edge of American technology has been blunted by the availability of research information through the U.S. Freedom of Information Act. Others attribute U.S. information loss to incoherent government information policies. It is difficult to assess the definitive effects of government policies on the growth and development of the information industry. However, it appears that nations like Japan, which promote and nurture infant industries, are highly successful in the manufacture and sale of sophisticated equipment used in information processing. This seems to be the case with respect to computers, photocopying devices, telefacsimile machines, and optical and video disk technologies. A diminished U.S. competitive position in the consumer electronics industry has been attributed both to protectionism overseas and to a decline in technological innovation in the production process.

While it has lost ground in the global computer and electronics industries, the U.S. still retains a large-scale capacity for the production of online databases and the creation of software used for the effective storage and retrieval of information. In 1987 nearly five hundred new online databases were created, bringing the total number available to approximately four thousand. Improved networking and marketing have led to the development of companies that specialize in providing access to groups of the most important databases; those concerned with science, technology, medicine, and law are among the most heavily used.

The potential beneficial and harmful effects of the transfer of information technology on developing countries have been studied recently by Shields and Servaes.[14] They concluded that the introduction of modern information technology can be helpful in the developmental process if it is properly balanced between self-reliance and strategic importation. Successful stimulation of economic and social development requires governmental political support and involvement of the intended beneficiaries in the planning and implementation stages. Involvement of potential information system users in the selection process tends to reduce the probability of mismatches between technology and information needs.

---

[14] Peter Shields and Jan Servaes, "The Impact of Information Technology on Development," The Information Society 6 (1989): 47-58.

F. A Proposal to Improve Information Services to Science and Engineering

To satisfy more fully their library-based information needs, scientists and engineers may require the creation of a National Library for Science and Engineering. Such a library might be based on the model of a central library such as the Canada Institute for Scientific and Technical Information or the National Library of Medicine, or it might involve the creation of a science library system through a consortium of existing strong science libraries. For example, a consortium of research libraries might be selected as regional centers on the basis of their collections, commitment to service, and willingness to cooperate. With a federal subsidy and a working board of directors representing these institutions and the scientists and engineers to be served, our nation could be expected to have at least one copy of all major print and nonprint materials--including databases--relevant to science and engineering in the consortium collection. With modern communication technology, data and information could be provided efficiently and economically to academic science centers, scientific associations, and industry as well as to individual scientists and engineers. The creation of an active science division within the Library of Congress would be extremely helpful in satisfying the literature requirements of science and engineering.

Management for connecting libraries should be supported by federal funds to establish and maintain the truly national nature of the enterprise. Support for special services such as interlibrary loan and special computer-based searches should be provided by the end-user who benefits from the service. Organizational problems associated with management may prove to be more difficult than the technical problems associated with networking operations and services. NSFNET, expanded to NREN, is a logical base upon which the network connections can be made. The scientific and engineering communities have a large stake in the development of such a library system and should be involved in the planning stages.

Uses of Science Information Indicators

Indicators that relate the communications infrastructure to the research enterprise may be used for several purposes. Kruytbosch and Burton have described how science indicators have successfully monitored the achievement of intrinsic goals of science, combined with political initiatives that seek to improve the performance of science.[15] Kochen has suggested that information systems may be developed to assist in science planning and policy making based on a detailed analysis of specific indicators associated with their monitoring and use.[16] In a general way, knowing how the entire communications infrastructure serves science and engineering allows the managers of science programs to seek ways to improve the efficiency and scope of the services provided.

Potential New Indicator of Science and Engineering

It has been suggested by Coyne[17] that a descriptive model of world science can be built from reference citations contained in the Science and Social Science Reference Indexes. Tapes containing 700,000 citations from 5,000 scientific journals can be searched and analyzed after the data are grouped into 350 research categories composed of 37,000 research specialties. Coyne believes that this analytic process provides a useful tool "for strategic planning and related management purposes." The National Science Foundation should consider testing Coyne's model to see if it could supplement the citation indexing currently used in Science and Engineering Indicators.

---

[15] Carlos Kruytbosch and Lawrence Burton, "The Search for Impact Indicators," Knowledge: Creation, Diffusion, Utilization 9, no. 2 (1987): 168-72.

[16] Manfred Kochen, "Models of Scientific Output," in Toward a Metric of Science: The Advent of Science Indicators, ed. Yehuda Elkana et al. (New York: John Wiley & Sons, 1978), 131.

[17] Joseph T. Coyne, "Descriptive Model of World Science," Information Hotline (December 1989): 9.

Information systems that serve science may also be used to inform the public of major developments and--through special education programs--improve the public's understanding of science. The mass media play an important role in this regard; thus, information and communications systems that serve science should also be available for use by others interested in science. In this context, the Congress, the Office of Management and Budget, and the major science agencies of government may be better informed about trends in science and engineering as they may affect public policy.

Communications in Support of Science and Engineering

Section III

TRANSACTIONS OF THE CLR/NSF CONFERENCE

October 30-31, 1989

Washington, D.C.

Eleanor W. Sacks

Table of Contents

Introduction: The Enterprise III-2

The Information Infrastructure

The Current Situation of Science Libraries III-4

Scientific Publishing Today III-11

Computers and Networks III-13

The Effect of New Technologies on Users

Changing Information Requirements of Scientists and Engineers III-16

Plans and Current Efforts to Meet Changing Requirements III-19

Organizational and Economic Implications of Technological Changes III-21

Recommendations for Further Study III-26

Introduction: The Enterprise

The Council on Library Resources (CLR) was asked by the National Science Foundation (NSF) to identify and analyze the major trends in the acquisition, use, storage, and dissemination of scientific and engineering information and to provide an analytical framework for use in future issues of Science and Engineering Indicators. At present, Science and Engineering Indicators is used by the executive branch and Congress for policy development and influencing resource allocation at the national level.

The Council brought together a small working group of key people in the areas of science administration, computing and networking, and science and public policy to meet with the program staff of CLR and discuss the directions and implications of the information revolution brought about over the last two decades by advances in computing and communications technology. From this meeting the Council developed the agenda for a two-day conference on "Communications for Support of Science and Engineering." For this meeting, held in October 1989, the Council assembled a panel of professionals involved in a wide range of research, administration, and policy-making activities. The panel included presidents and senior administrators of universities with strong science and engineering programs, directors of university libraries specializing in science and engineering, a scientific journal editor, and social scientists interested in the impact of new technologies on human information-seeking behavior. The group met with CLR program staff and senior NSF personnel from the offices of networking and communications research, science indicators, and special analytical studies. A list of conference participants is found at the end of this appendix.

The first morning session included three presentations on the history and current state of the information infrastructure serving science and engineering as represented by libraries, scientific publications, and computing and networks. The afternoon session examined the changing information requirements of scientists and engineers, plans and current efforts for meeting these changing requirements, and the organizational and economic implications of technological changes. With a base for discussion established, the second day was devoted to an examination of the relationship between information services and scientific activity, to identifying new indicators that might be used to measure the progress and productivity of science and engineering, and to the studies that need to be performed to evaluate more fully the information resources supporting science and engineering. The recent research undertaken by the National Library of Medicine (NLM) on the impact of its MEDLINE database on patient care was presented as a case study.

I. THE INFORMATION INFRASTRUCTURE

A. The Current Situation of Science Libraries

What are they? What do they contain? What are the costs? What do they do? These are some of the questions addressed by Patricia Swanson in her report on the current state of science libraries in the United States. Libraries in support of science number in the thousands and come in many forms: major research libraries that are either university-based like MIT or public like the New York Public Library; college and university libraries; medical and science libraries; special libraries serving corporations and quasi-independent research institutes like Woods Hole and Fermilab; and federal libraries, including the Library of Congress and those in the national laboratories. All of these libraries support the range of scientific research from theory to commercial applications.

Today the collections of libraries are still primarily on the printed page in the form of books and journals. But libraries also contain much more: maps, specifications, patents, "semi-published" literature such as working papers, technical reports, preprints, and, increasingly, information in data and electronic formats--magnetic tape, floppy disks, or optical disks. The amount of material available to scientific researchers is staggering. Of the 260,000 active serials worldwide listed in the International Serials Data System, approximately 50,000 titles are related to science research. Chemical Abstracts indexes about 9,000 titles (450,000 citations each year); Index Medicus covers 2,888 publications (250,000 citations per year); 520,000 citations were added to the BIOSIS database in 1988. The National Technical Information Service (NTIS) adds to its list approximately 70,000 technical reports per year.

Even as the scientific literature has expanded at an explosive rate, the ability of libraries to acquire the increased volume of information has declined severely because of drastic increases in costs, primarily of journal subscriptions. An example from the University of Tennessee characterizes the national dilemma. Over the past twelve years a 70 percent increase in funds spent on serials resulted in a 30 percent decrease in the number of titles acquired. The growth in the size and cost of the serial literature threatens the scientific enterprise not only because libraries are unable to maintain and expand their collections, but also because of the vast implications of that growth for the dissemination of scientific information.

The amount libraries spend in support of science is substantial. Statistics from the over 100 members of the Association of Research Libraries for 1987-88 showed that nearly 435 million dollars were spent on all forms of library materials. If even one-third of that amount were in support of science, the ARL libraries alone would have spent nearly 145 million dollars that year.

Libraries provide three basic services for their scientific clientele: they discover, locate, and deliver information. While these services have been constant, the techniques, strategies, and systems underlying them have changed significantly because of changes in technology over the last fifteen years. Librarians have been moving from fulfilling these functions from their own collections to increased reliance on the collections of other libraries and the services of other information providers. The trend in library functions has been shifting from ownership of information to access to information.

The "discover" function once meant reading current journals, searching the locally held collection, or guessing which institution might contain needed information not available on-site. Today it means online catalogs available over campus networks, backed up by national electronic bibliographic utilities, such as OCLC, Inc. (3000 contributing libraries); the Research Library Information Network (RLIN--36 members); and a host of electronic, external databases. The local and extended knowledge universe is now being integrated into a single system.

In libraries, the greatest revolution created by the introduction of the computer has been the expanded capacity for discovering existing literature quickly and easily. However, this dramatic progress has heightened the pressures on librarians to locate and deliver information. What a scientist did not know existed two years ago must now be delivered in a few hours. As the "visible universe" of literature and data has expanded, important information is coming from unexpected and previously unknown sources.

The expectations created by this enhanced ability to find information are great. In the future, libraries will require that more data--both bibliographic and text--be available in machine-readable form. This will be accomplished by the retrospective conversion of library catalogs into electronic formats. In addition, libraries will have to insure the survival of existing information through preservation efforts.

Procuring information as a scientist needs it, rather than buying it in anticipation of need, may become an increasingly reliable and economical strategy for meeting information needs. This strategy has been made possible by technology in the form of national scholarly computer networks and local campus networks. Delivering information has been revolutionized in many libraries; delivery now takes the form of electronic requests, campus delivery to offices and labs, electronic file transfer, and telefacsimile. Electronic networks such as OCLC, RLIN, and NLM's DOCLINE, in addition to identifying and locating materials, offer national systems for interlibrary lending. Traffic on these systems demonstrates their value. In the seven months from October 1988 through April 1989, 927,688 health sciences information requests (an increase of over 100,000 from the same period in the previous year) were input into DOCLINE by over 1700 participating libraries. (In April of 1989 the daily input averaged 6,729 requests.)

Despite this shift in the way science libraries can deliver information, individual libraries differ considerably in performance of their stated mission to acquire and preserve material locally. Large research libraries remain the most traditional in their focus on local collecting and permanent ownership as the best long-term means of identifying, locating, and delivering materials to their own patrons and fulfilling their function as external resources for other science libraries. Corporate libraries often collect only to meet their current needs, discarding materials when researchers move on to other areas, and rely very heavily on external sources such as other libraries and commercial information brokers to respond to individual requests. The degree to which libraries have reallocated funds from local purchasing to access is less than revolutionary, however, with the greatest reallocations seen in smaller academic institutions with curriculum emphasis on applied science and in special libraries.

The increased emphasis on delivery by drawing on resources located elsewhere requires libraries that are organized to interact with other institutions and information sources. They must be able to create internal technology for external access (e.g., electronic bibliographic records that can be shared, local systems linked to national databases or periodical indexes, and library-to-library connections through national electronic networks).

There are two patterns of external organization demonstrated by science libraries in the United States today. The first is the organization of the health science libraries created by the Medical Libraries Act of 1965, which initiated a top-down, hierarchical structure of designated regional resource libraries capped by the National Library of Medicine. NLM, with its mandate for comprehensive collecting, stands as the last-resort library for information seekers in this system. But more importantly, it functions as the impetus for technological developments that began with the MEDLARS system for bibliographic citations, extended to other specialized databases, and are now moving beyond citation databases to sophisticated systems serving specialized areas. Some examples are systems that supply toxicology information, provide diagnostic support, or supply full text of needed documents. NLM developed, implemented, and continues to financially support DOCLINE, the automated interlibrary loan network. In addition, NLM has fostered the application of information technologies through its Integrated Academic Information Management Systems (IAIMS) programs.

Unlike medical libraries, science libraries in the United States do not have a hierarchical structure topped by a library that is clearly dedicated to being the source of last resort for publications or the initiator of new information technology, such as the National Library of Medicine. The Library of Congress pioneered the MARC (MAchine-Readable Catalog) format for providing bibliographic data that is the backbone of automated bibliographic efforts. The National Agricultural Library produces the major database in the field of agriculture and is developing an agricultural library network. The role of these libraries as national resources is less clearly defined than NLM's. Other libraries supporting science interact more or less informally, using facilities such as OCLC and RLIN and by being a part of regional, state, or metropolitan networks or federal consortia such as the FLICC or the CENDI group. Whether a more defined national science library on the model of NLM is needed is a topic worthy of full discussion. The creation of such a library through a consortium of existing science libraries deserves study.

Examination of current library operations shows that as libraries have moved from ownership to access, the information is still delivered primarily in paper formats. For the access model to be truly effective, it must be able to provide information in a timely fashion. That will require a major effort to transform print into electronic formats. But as the movement to access takes place, it should be remembered that some institution must own the information before others can have access to it. The question arises: Which organizations should assume stewardship responsibilities over time for the resources to be delivered on demand? And which organizations shall reap the benefits of ownership, if any, or pay its price? Technology is causing a revolution in libraries, but the process is slow and the transition is still under way.

B. Scientific Publishing Today

Philip Abelson reviewed the history of scientific journal publishing in order to elucidate how the current publishing structure was created and how it has evolved. Scientific publications began as the products of scientific societies. The societies saw the publication and dissemination of scientific results as being for the good of the profession, and their journals were produced for the most part with donated labor. As the demand increased for more and more publication space, the societies found that the costs of launching new journals were prohibitively expensive. The societies developed two strategies to help defray costs: they instituted page charges that were assessed on the authors and started charging differential fees for journal subscriptions, requiring libraries to pay two to five times more for subscriptions than members of the societies. The scientific societies were slow to initiate new journals, especially in multidisciplinary fields. In this economic situation, commercial publishing companies, mostly foreign-owned, saw an opportunity to expand their scientific publishing by identifying and producing journals in these new multidisciplinary areas. The publishers attracted authors by eliminating page charges and appointing them to editorial boards. When publication was under way, the commercial companies enhanced their revenues by further increasing the fees charged to libraries. The production of scientific journals proved to be so profitable that other commercial companies moved in, with the result that more than half of scientific publishing in the U.S. is commercial and a large percentage is foreign-owned. The rising costs of journal subscriptions are painful and the increases are rising far faster than inflation. Commercial firms are now moving into electronic publishing as well. There is some hazard in allowing commercial firms to control the majority of scientific publishing, since each journal has a de facto monopoly on the information contained. The great increase in journal costs has meant less access to information by professionals. At the moment the interests of the publishers appear to outweigh the interests of the authors and the readership to ready access, according to Dr. Abelson.

In further remarks Dr. Abelson stressed that human interaction is what leads to creativity. Technology will not fundamentally alter that fact, although over the years it has greatly facilitated human interaction through ease of travel and advances in telecommunications. An important trend over the last decade that assists increased human contact has been the rise in the number of closed symposia and meetings where researchers can meet face to face.

Dr. Abelson noted that, thus far, electronic databases have had a differential impact on disciplines and even within subdisciplines. In biology, for example, the Genome Project is the biggest database, but many biologists have little involvement with, or interest in, electronic databases. In the arena of biotechnology and pharmaceuticals, where there is a deep interest in commercial applications and patents, the transfer of information is handled quite differently. While information in electronic formats is increasingly important, traditional publishing will continue to have great utility, and Dr. Abelson feels it should continue to have priority.

C. Computers and Networks

William Arms reviewed network development and utilization. The past decade has seen spectacular changes in computing brought about by the emergence of the personal computer. Equipment is now smaller, cheaper, and faster, and the economies of scale have largely disappeared. The big central campus computer has receded in importance to become only one node on a large network of computers, and the network has become the central feature of the computer system. In universities, most faculty and research staff have personal computers, reflecting major advances in user interfaces and user friendly software. A survey of faculty at Carnegie-Mellon University showed that personal computers are ubiquitous: 92% have computers; 90% have computers in the office; 76% said they use them at least one hour per day; 43% have a UNIX or VMS workstation; 67% have a large personal computer; 17% have laser printers at home; 53% have a network connection in the office; only 19% use central computers on a regular basis (much of that use is for electronic mail); 10% use supercomputers -- 7% the local NSF-sponsored supercomputer center and 3-4% supercomputers elsewhere, some as far away as CERN. The ways researchers use computers break down as follows: 90% do word processing; approximately 75% use computers for electronic mail; the same percentage use online library services. Fifty-six percent have more than 10 years computer experience, and 27% consider themselves expert programmers. The planning assumptions for the five-year period from 1990 to 1995 at CMU are that every scholar has a powerful personal computer with a 1,000 by 1,000 pixel display, with all computers connected to a network by links of 1.5 Mbit/sec or faster. The forecast is that by 1995 these assumptions will be seen as restraints rather than the norm.

The means for transmission of electronic information has taken several new forms. The first is electronic mail and computer bulletin boards. Vast quantities of scientific information are now reaching individual researchers through these avenues well before journal publication. Telefacsimile transmission has also grown in importance, especially for international correspondence. A good estimate is that the rate of transmission or amount transmitted doubles every three months. Third, there has been a steady build-up of information online, as evidenced by the fact that full-text databases have now entered the mainstream.

Information itself has also taken new forms. These include dynamic databases that are updated daily, such as those being developed at the Welch Library of Johns Hopkins; electronic documents, such as mathematics papers that can be used interactively by students to manipulate the mathematical formulas;

hypertext; and an increasing emphasis on mixed text, video, and computing. A new area that libraries need to be involved in is collecting information that is not print based, especially video. For example, it is hard to understand recent politics without being able to view tapes of political speeches and campaign commercials.

Given the advances in computing and networking, it is now possible to conceive of an electronic library. The electronic library would have the capability to access and deliver information from anywhere in the network to where the information is needed, including faculty offices and laboratories. Computer power is used for searching and retrieval and to simplify identifying the location of information. The information would be kept current because it would be periodically updated. The electronic library may have the long-term potential for being cheaper than the traditional library.

Access to computers and computing systems is now essential to the conduct of scientific research. In the world of computer science, where access to networks and bulletin boards is open and relatively easy, the free flow of information has had a democratizing effect; researchers can work with anyone almost anywhere in the western world. In other fields where access to specialized research networks is expensive and controlled by the elite researchers with large research grants, the effect has been the opposite, reinforcing control by the powerful. The issue of access to information is a matter for full public consideration that should be addressed by a national body such as the Office of Science and Technology Policy. But it is also a problem for campuswide attention. Too often, general-purpose databases such as ERIC are purchased by individual departments and schools rather than by a central resource like the library, and, as a general rule, the entity that pays for the information controls access to it.

A further hindrance to research information for many scientists is their own lack of expertise in making full use of computer systems. Many of the specialized networks suffer from a lack of consistent protocols and finding aids. There is a great need for standards in computer interfaces. Without computer professionals who can decipher the differences, the difficulty of using these networks effectively blocks access to information on what is happening in the wider research community.

II. THE EFFECT OF NEW TECHNOLOGIES ON USERS

A. Changing Information Requirements of Scientists and Engineers

Dr. Donald Langenberg chaired a panel of the National Academy of Sciences (NAS) that produced the report Information Technology and the Conduct of Research: The User's View (1989). The panel's mandate was to study how scientists use information technology. Little is known about how scientists and engineers use information; literature in this area is scarce, and the information the panel collected was for the most part anecdotal.

The subject of how scientists and engineers use information is a fruitful area for research investigation.

The NAS panel put together a general picture from which it is clear that information technology has already had a substantial effect on the conduct of science and engineering. In most fields the history is similar. Computers started as number crunchers, doing calculations bigger, faster, and better than the humans and mechanical equipment they replaced. The use of computers changed with the development of the ARPANET, the first national computer network, which was funded by the Department of Defense. The DOD saw this network as a way for researchers to share resources, primarily for number crunching. However, the ARPANET had benefits that the Defense Advanced Research Projects Agency did not anticipate. Most importantly, it facilitated informal communication and the sharing of scientific information, not just the sharing of computer resources.

The advent of smaller, more powerful computers--minicomputers, workstations, and PCs and their equivalents--had a major effect on the conduct of research, bringing large computing capacity into individual research labs and faculty and student offices and making access to networks even easier. Now research has developed to include projects that could never be done without computers. Two examples are unmanned space exploration and high energy physics, in which computers control experiments and equipment that could not otherwise have been designed, built, and operated.

Computers have had a major impact on the generation, as well as on the acquisition and storage, of data. The Library of Congress contains 10 terabits of information gathered over two centuries. Medical scanners produce that much information every week, and the Superconducting Supercollider scanners produce that much information every second. At present, dealing with such quantities of information is impossible, with or without computers.

The research community both uses and produces information technology. For example, a very large scientific computing capacity is now needed to design the next-stage supersonic airplane and for massive storage of information. The bigger problem at present, however, is not hardware but software. Hardware development is advancing at a far faster rate than the software needed for effective operation. Currently, researchers face the following software options: rely on off-the-shelf software that may not meet their needs; design their own that may be too customized and not fully documented for others to use; or do without.

To describe the components of the new information structure in its totality, Dr. Langenberg has coined the term "infory" to include the mail system, telephones, fax, computing center (increasingly the unit that manages the campus network), and the library. Infory provides (1) places to store and communicate information; (2) communications--across the campus and around the world; and (3) people who can help--from mail handlers and telephone operators to computer specialists and librarians.

B. Plans and Current Efforts to Meet Changing Requirements

Kenneth King spoke on "The Impact of Technology on Information Needs in Science and Engineering." If, as is commonly stated, it takes fifty years for a technology to advance to the point where it fundamentally alters the lives of almost everyone, the age of computers and communications, begun in the 1940s, should be fully in place during the 1990s. Advances in computing power are proceeding at an explosive pace. It takes 1 Mips (Million instructions per second) to control words and 100 Mips to control pictures and sound. It will take 1000 Mips for voice recognition, machine vision, and language translation, and 10^{12} Mips to simulate physical systems directly from the fundamental equations of physics. In higher education today, four important new capabilities are emerging: first, the ability of scholars to communicate with other scholars electronically and access distributed knowledge databases and experimental instruments through high-speed international networks; second, the ability to build knowledge on a new and dynamic electronic platform consisting of text, numbers, images, and sound; third, the ability to build complex software systems from components; and fourth, the ability to study the properties of complex systems by creating information analogues of these systems in computers. Out of this is emerging a new structure on which to build knowledge. This structure will have the ability to represent, display, modify, and interact with information in multi-dimensional and multi-media formats that will include numbers, words, images, and sound. The interaction with this information will also be characterized by modular program elements that can be read, edited, and plugged in to create new objects. This ability will change the paradigm on which knowledge is built from a static one (print) to a dynamic one.

The twentieth century has seen the development of a third kind of science made possible by the power of computers. Science is moving from the realms of direct experimentation and theory to the new field of simulation through which scientists will have the ability to create information analogues of highly complex physical systems such as the dynamic patterns of weather.

The goals of networking in science and engineering are to connect every scholar in the world to every other scholar and thus reduce the barriers to scholarly interaction; to connect to the network all important information sources, specialized instruments, and computing resources worth sharing; to build databases that are collaboratively and dynamically maintained and that contain all that is known on a particular subject; and to create a knowledge management system on the network that will enable scholars to navigate through these resources in a standard, intuitive, and consistent way.

C. Organizational and Economic Implications of Technological Changes

Professor Robert Hayes of the University of California, Los Angeles, addressed his remarks to the management of information resources, focusing for the most part on the university environment. As outlined earlier, the information resources on campus include not just the library, but telecommunications, campus computing facilities, other repositories of nonprint materials such as archives of film, and specialized research tools such as large data sets.

Several management models have been developed in academia to handle the expanding and increasingly diverse forms of information. Some campuses have instituted a centralized administrative office or information tsar, whose mandate is to oversee and coordinate all of these areas. UCLA, however, has chosen a more distributed management model based on the specific concerns of its own institutional structure. These include the different kinds of expertise required to manage the variety of information resources, the prior existence of large bureaucracies, the fact that some resources are tied to specific programs, and the need to maintain separate cost centers for the various operations. But when considering the total amount of institutional financial resources devoted to the wide range of information resources, all of these areas must be combined. On the UCLA campus, approximately 10 to 15 percent of the total university budget is devoted to information resources. The total cost of the current 2,500 UCLA grants and contracts is approximately 10 percent of the budget.

Through a large multi-year grant from the Council on Library Resources to study "Strategic Management for Academic Research Libraries," Professor Hayes was able to identify examples of the current information resource needs of UCLA's academic programs. These specific needs coalesced into fifteen broad areas of concern that should have widespread applicability for the nation's research community.

1. The need for print publication continues.

Print continues to be an essential medium for scholarship and scholarly communication in spite of the advent of new information technologies. While some experts believe that technology could lead us to the "paperless society," others feel it could just as easily drown us by its ability to create a flood of new paper-based information.

2. Libraries/computers/telecommunications need to be integrated.

Integration does not necessarily imply a unified administrative effort, but it does mean that their parallel development must be closely coordinated.

3. Interinstitutional cooperation must be greatly enhanced.

Much academic research today involves cooperation among many institutions. Interinstitutional cooperation implies commitment of resources and, to a significant degree, loss of independence.

4. The problem of acquiring foreign materials needs to be addressed.

The acquisition of information materials from foreign countries is an especially acute problem with respect to the world's "developing areas." Our knowledge of these large populations and vast geographic areas is at best fragmentary. Coordination and cooperation between the federal and university libraries for acquiring such materials are necessary, and networking will play an enormous role in identifying and delivering the information available.

5. Library services need to be enhanced.

Two areas identified under this rubric are the need for current contents services and an inventory of campus databases. A project is under way at UCLA to experiment with a current contents and journal delivery service that is completely supported by subscriber members' fees. The UCLA library is attempting to inventory all of the databases that are available on campus--those that are formally acquired, those that are associated with large-scale projects, and those that are maintained by individual faculty members. The objective is to establish a basis on which catalog entries for databases can be incorporated into the online public access catalog.

6. Information centers need to be established.

There is a recurring need for information centers that will acquire information resources for specific academic programs that cross disciplinary lines. One example at UCLA is the Hazardous Substances Information Center.

7. User interfaces are essential tools in the ability to implement new technology.

Without effective user interfaces, access to information resources will be hindered and delayed, if not denied.

8. Digitized images present enormous problems of storage due to the magnitude of the files produced.

The problem is especially acute for those files arising from planetary and medical scanners. The primary concern is with the management of the files themselves. Fundamental research needs to be done on the organization of such files and on the means for retrieval from them.

9. Off-campus users present special problems.

Attention needs to be paid to the special needs of remote users, who will not have finding aids readily available to them.

10. Expert system development has broad applications.

In the field of library science, expert systems are being developed to support ready reference in academic libraries.

11. Indexing and abstracting services need to be extended.

Indexing and abstracting systems are essential support tools for information access, whether in printed form or through online access. These tools are well developed in the sciences and engineering, law, and medicine, but they are far less well developed in the humanities and the arts.

12. Desktop publishing is an essential tool for getting material into distribution quickly and maintaining control of the end product.

13. Database development by faculty will require expert assistance.

Faculty are developing databases across campus and most will need consulting support for design of database structures, indexing data files, and downloading data from external sources.

14. The need for file conversion is widespread.

As the research community moves from paper-based data to computerized data, the need to convert existing data sources to digital text or digital image grows. Optical character recognition equipment is well enough developed that it can now be expected that this application will be instituted on a large scale, but the biggest question is how to manage the process.

15. The need for project management support tools and consultative services grows as other fields outside the physical sciences and engineering take on large-scale collaborative projects.

Recommendations for Further Study

The conference participants reached a consensus on several broad areas for further study.

First, the current science indicators, especially those that measure the production of information in print form, should be continued. In conjunction with these indicators, the other ways scientists and engineers receive, access, and distribute information should be identified and measured in order to construct a fuller picture of the information infrastructure supporting science and engineering.

Second, the information-seeking behavior of scientists and engineers should be studied in order to better understand the information infrastructure in the United States today. It would be helpful to understand how they discover, select, and use the vast amount of information available to them. Much anecdotal information has been introduced to support the idea that easy access to computer networks and fax technology has made journals and even conferences obsolete as the primary modes of information transfer by facilitating the transfer of knowledge. Studying the information behavior of researchers will test the truth of these observations and delineate the distinct patterns among the different branches of science and engineering. Information technology has changed the ways information is stored and accessed; the need to look closely at the changes taking place and their implications is a major concern.

Third, if the idea that easy access to information has a major impact on the strength and productivity of U.S. science and engineering is accepted, studies are needed of a number of perceived barriers to access. Some of these barriers are: the high costs of published and electronic information; the unequal distribution of information technology across disciplines and between the corporate and academic worlds; the lack of standards for computer interfaces that constrain ready access; the effect of patent and copyright regulation on information flow; and an ambiguous national information policy.

Fourth, the management of information resources requires attention. The moment is fast approaching when it will not be possible to manage the information generated by information technology. Strategies must be developed for projecting the size of new data sets and dealing with the immense problems of management and storage that they create.

In the field of medicine, the National Library of Medicine has taken the lead in organizing the information infrastructure and adapting information technologies to distribute that information widely to researchers and practitioners. In science and engineering, however, no such national capacity has emerged. As a result, there are vast differences in information access among disciplines. The need for the creation of a National Library for Science and Engineering deserves study. In computer science, development of the ARPANET (the first international computer network devoted to research) was supported by the Department of Defense, and communications among researchers have been characterized by open and easy access to information. Other disciplines have benefited from the advances made in computer networking with the expansion and multiplication of specialized computer networks, but access to pertinent information is not assured.

Specific Topics and Possible Projects

1. New Measures of Information Resources
    a. Provide an overview of science and engineering information resources that are not print-based;
    b. Compare and correlate the growth rates of print-based information to the growth rates of electronic information;
    c. Collect data on the size and type of science collections in major research university libraries and seek ways to correlate that data with the quality and productivity of graduate programs at those institutions;

d. Survey the numbers and types of databases available commercially and within the federal and academic research communities;

e. Collect data on the growth and use of information technologies--electronic mail, fax, alternate forms of publishing (desktop and electronic);

f. Survey trends in libraries supporting science and engineering--acquisitions, storage, access, delivery;

g. Collect and compare data on the financial resources allocated/expended for information resources in university budgets, academic research contracts and grants, and corporate research and development.

2. The Information Behavior and Requirements of Scientists and Engineers

a. Compare the information-seeking behavior of younger scientists with more established scientists;

b. Compare the science and engineering libraries in support of corporate research to those in academia in order to ascertain the differences in the ways the two realms access information, how scientists and engineers differ in the ways they use information, and what the impact of different information-seeking patterns is on their research;

c. Study the trends in scientific and engineering conferences as an indicator of the growth and strength of new fields and the lessened activity in other fields. As new fields are in the process of being created, the first step is often a conference devoted to defining the new area. As fields expand and mature, conferences often subdivide into more refined groupings. As areas recede in importance, the numbers of conferences decline and may even disappear.

3. Access to Information

   a. Study ownership and control of access to information;

   b. Consider the need for consistent user interfaces and the establishment of national standards to advance easy and equitable access to information.

4. Management of Information Resources

   a. Study the problems created by the explosive growth of information resources, especially the growth of massive data files produced by medical and physical scanners, the implications for storage, the need for coordination among research groups, and the problems of accessing such vast amounts of data;

   b. Explore the need for repositories of data sets;

   c. Examine the role of the existing federal libraries (or creation of a National Science Library) in advancing information technology and in organizing the information infrastructure.

Participants in CLR/NSF Conference
October 30-31, 1989

Philip Abelson
AAAS (Science)

William Arms
Carnegie Mellon University

William Baker
CLR Board

Robert Hayes
University of California, Los Angeles

Sara Kiesler
Carnegie Mellon University

Kenneth King
EDUCOM

Donald Langenberg
University of Illinois at Chicago

Peter Likins
Lehigh University

Jay Lucker
Massachusetts Institute of Technology

Nina Matheson
Johns Hopkins University

Thomas Shaughnessy
University of Minnesota

Elliot Siegel
National Library of Medicine

Patricia Swanson
University of Chicago

Duane Webster
Association of Research Libraries

CLR staff:
Martin Cummings
Warren J. Haas
Henry Riecken
Eleanor Sacks

NSF staff:
Lawrence Burton
Jane Caviness
Darleen Fisher
Carlos Kruytbosch
Leonard Lederman
Rolf Piekarz

Section IV

SPECIAL STUDIES RESULTING FROM THE CONFERENCE

CLR asked Helen H. Gee, former Chief of the Office of Program Evaluation, National Institutes of Health, to examine the current information-seeking practices of research scientists. Dr. Gee prepared a working paper that served as the basis for discussions with a panel of experts in scientific communications. Her final report, which was informed by these discussions, (1) addresses the role of new information technologies that influence information-seeking practices and the relationship of these processes to research activity and productivity and (2) describes the types of studies and research strategies recommended for use in planning for improved information services for science and engineering.

The Council also commissioned a study of the relationship between productivity in academic and engineering research and the extent and character of library and information resources available to support the research enterprise involved. The principal investigator, Nancy Van House of the University of California, Berkeley, assessed the quality of scientific research in relationship to faculty membership in organizations such as the National Academy of Sciences, the National Academy of Engineering, NSF Presidential Young Investigators, Nobel Prize winners, and other comparable groups. Universities with significant science activities provided data to measure library resources and services utilized by scientists. An effort was made to correlate faculty quality with ranking in library resources. Differences between outstanding research universities and other broader-based universities were analyzed in an attempt to explore the possibility of a relationship between resources and research accomplishments.

David Penniman, Director, Libraries and Information Systems, AT&T Bell Laboratories, and Jay Lucker, Director of Libraries, Massachusetts Institute of Technology, have undertaken a study to compare costs of library and information services in support of science and engineering performed in academic and industrial institutions. They expect to complete the study in the fall of 1990.

THE USERS AND USES OF SCIENTIFIC INFORMATION RESOURCES:

Recommendations for Study

Prepared for the Council on Library Resources

by

Helen Hofer Gee, Ph.D.

May 1990

The Users and Uses of Scientific Information Resources: Recommendations for Study

INTRODUCTION

How are present-day scientists actually using the countless information and communications resources that are now available? What roles do these tools, technologies, and services play in the process of planning, conducting, and reporting on scientific research? To what extent do availability and access to such resources influence and affect productivity? And how seriously disadvantaged are scientists whose access to information technology is limited, when at the same time access to traditional journals is being severely curtailed? These are a few of the questions that are troubling a growing number of leaders in science, education, administration, and government. Answers are needed if sound decisions about the allocation of funds in support of research are to be made during the next several years.

The Council on Library Resources (CLR), in October 1989, convened a meeting of senior administrators, including directors of large science libraries, computer and information scientists, and leaders in the science and public policy community to discuss the direction and implications of advances in computer-based communications technology. CLR had been asked by the National

Science foundation (NSF) to "identify and analyze the major trends in the use, storage, and dissemination of scientific and engineering knowledge and to provide an analytic framework for use in future Science and Engineering Indicators volumes that would be of interest to and useful for the various branches of the federal government and the science and engineering communities at large."[1]

The CLR group reviewed the enormous growth in the quantity and format varieties of information that the scientific community is producing, and to which it requires rapid and efficient access. The technological advances that have so greatly increased the ability to discover the existence of new information were recognized as also having resulted in enormous increases in the demand for access. At the same time, crippling increases in costs of scientific journals (due in no small measure to increasing commercial ownership) were identified as a significant force for changing how libraries, especially university libraries, must function in order to meet their obligations for access.

Electronic interlibrary networks, file transferring capabilities, and telefacsimile have enlarged libraries' capacity to share each other's resources and to procure information as needed rather than acquiring it in anticipation of need.

---

[1] Report on the CLR/NSF Conference on "Communications for Support of Science and Engineering," October 30-31, 1989 (Draft, p. 7).

Increased emphasis on the delivery of library materials that are located elsewhere also requires reconsideration of the effectiveness of the external organization of U.S. science libraries, as well as the adequacy of local and national computer-based systems, databases, indexes, and networks. Also recognized was a need to study the information-seeking behavior of scientists and engineers to understand how they discover, select, and use the vast amount of information available to them.

Consideration by CLR of ways to act on conference results has led to preparation of this paper. To that end, experts in psychological measurement, communications, and physical and information sciences research were convened by CLR on February 27, 1990, to discuss possible approaches to studying the uses of information resources and services by scientists.[2] The need for further development of the Science and Engineering Indicators' analytical framework relating to national information resources served as a general background for these discussions.

Three possible approaches to improving current knowledge and understanding of how scientists use information resources were carefully considered:

---

[2] Participants in the meeting included: Porter Coggeshall, National Academy of Sciences; Robert Kraut, Bell Communications Research; Carlos Kruytbosch, National Science Foundation; Leah Lievrouw, Department of Communication, Rutgers University; David Penniman, AT&T Bell Laboratories; Jeffrey Mandula, Division of High Energy Physics, Department of Energy; and Michael Rappa, Sloan School of Management, Massachusetts Institute of Technology.

* Replication of studies conducted primarily during the 1960s to update findings. In these studies, the dynamics of scientists' uses of information resources were studied in relation to variation among disciplines, types of research conducted, career stage, etc.

* Conduct of one or more broad-based, multi-institutional, multidisciplinary surveys of a large variety of scientists' current uses of and needs for both conventional and technologically advanced information resources.

* Support of research programs designed to encourage intensive experimental, laboratory, and critical incidents approaches to understanding how advancing information technologies are influencing and changing patterns of behavior, and the needs and opportunities scientists have for obtaining information resources.

BACKGROUND AND ALTERNATIVES

Understanding how information resources in general, and advancing technologies in particular, affect the activities and performance of scientists is an exceedingly complex enterprise. Fundamental descriptive data are sparse and mostly anecdotal. We know far too little about how information resources are actually used by scientists as they plan, design, conduct, analyze, and report on their research.

During the 1960s, when the computer and communication sciences were in the early stages of revolutionizing the conduct of scientific research, William Garvey and his colleagues at Johns Hopkins University conducted an extensive series of pioneering surveys and social psychological studies of communication among scientists. A "scientific information crisis" was already believed to exist in 1961. Garfield and Griffith noted their impression that, in communication among psychologists, the components of the system of communication seemed to compete with one another rather than contribute a particular function, and that objectivity played no part in attempts to govern and revise the system.[3] These impressions led to a focused series of some seventy studies of the information-exchange activities of more than 12,000 scientists and engineers from nine physical, social, and engineering subdisciplines. The investigations included "the full spectrum of activities associated with the production, dissemination, and use of information from the time the scientist gets the idea for his research until information about the results of this research is accepted as a constituent of scientific knowledge."[4] The communication behaviors studied ranged from the most informal discussion of a pair of colleagues to formal publications.

---

[3] Garvey, William D., and Belver C. Griffith, "Scientific Communication as a Social System." In William D. Garvey, Communication: The Essence of Science, 149. Oxford: Pergamon Press, 1979.

[4] Garvey, William D. Communication: The Essence of Science. New York: Pergamon Press, 1979.

Garvey and his colleagues laid a foundation on which a great deal of subsequent communications research has been based.

When the Garvey studies were performed, nearly a quarter century ago, journals, books, local colleagues, and those reached by letter and telephone were the scientists' principal, and in many cases only, sources of information available for assistance. Scientists were heavily dependent on their own and their universities' library collections and catalogues for published information. In 1972, a National Academy of Sciences (NAS) committee reported on physical scientists' uses of information resources. The amount of time physicists spent using computer services as a communication medium was too small to gain even a blip on an "hours per week" graph. Chemists spent an average of thirty minutes using computer services, compared with four hours per week in oral communication with colleagues. Physicists spent an average of eight hours in talking and listening to colleagues.[5]

After completing their studies, Garvey and Griffith stated that scientists "...no longer view the crisis *simply* as an information flood, for now, after several years of extensive planning, developing, and trying out of national, discipline-oriented information systems, it is apparent that these systems, which promised much, have largely failed in terms of attracting widespread use...."[6] In spite of or perhaps because of this failure, the past two decades have seen such a proliferation of new communication and information technologies that few scientists outside of the computer and communications disciplines have been able even to keep themselves informed of the developments.

---

[5] National Academy of Sciences, NRC, COSEPUP, Physics Survey Committee, D. A. Bromley, Chairman. "Dissemination and Use of the Information of Physics," in *Physics in Perspective*, Vol. 1, Chapter 13, 1972.

Although there have been efforts to discover how information systems are used, and the nature and extent of their impact, very little generalizable information is available because most field research has been conducted in single organizations, using a single system in a unique organizational environment. Furthermore, investigative efforts have not been focused on issues, questions, or populations that might be useful at a national level. At best we have some useful comparisons of computer-mediated communication (CMC) systems against other media in local circumstances.

The fairly extensive 1960s studies of uses of information resources were limited to studying how groups of scientists from different disciplines (samples drawn from membership in professional societies) used printed and personal communications resources. These studies were not designed to investigate the extent to which significant situational determinants of behavior (such as organizational and institutional characteristics, sources and extent of research support, or availability and access to computer based technologies) may interact with other variables in influencing behavior patterns. Simply replicating earlier studies would not achieve better understanding of present-day information-seeking behavior, but coupling the results of a new survey with those of earlier surveys and behavioral studies could provide information about the kinds of changes that have occurred.

---

[6] Garvey, William D., and Belver C. Griffith, "Communication and Information Processing Within Scientific Disciplines: Empirical Findings for Psychology," in Garvey, William D. Communication: The Essence of Science, Appendix A, p. 127, 1979.

A suitable broad base of information on the scientific research community's acceptance and use of the new information technologies would help to guide research on their impact on the conduct of science, on the behavior of scientists, and on how, for example, productivity may be affected by changes in the systems through which resources are made available. Academic, professional, and government planners and decision makers clearly need such information. The information industry would also benefit from a broad survey of current knowledge, awareness, and actual utilization of information resources; the information could profitably improve their identification of future markets. The above considerations all point to the need for a comprehensive, national survey of present day scientists' and engineers' uses of information resources.

Of course, a single survey yields only one macroscopic view of current information users and their circumstances. Its results can suggest but not confirm how personal characteristics and situational determinants interact to produce the observed varieties of information-seeking behaviors and activities. The value of a current survey can only be fully realized if it is subsequently expanded by more focused research on these dynamic relationships. Plans should be made to repeat the survey within a period of about five years so that the nature and directions of change can be accurately assessed.

THE QUESTIONS

It is time to establish a new empirically determined base of knowledge about scientists' uses of information and communication resources. Such a knowledge base could serve as a springboard for more detailed investigation of the dynamics that underlie variations in patterns of utilization. It will be prudent and economical to conduct a broad survey of the entire academic scientific community within a short time span so that determinants and interrelationships can be identified without the complications of continuing change that an extended time span would introduce. While the focus of attention is primarily on the behavior of individual scientists, some questions also need to be addressed to institutions and their scientific subdivisions, and others to university libraries.

Some of the questions to which fairly comprehensive answers could be obtained from a single survey follow. The kinds of questions that would be answered by querying individual scientists and engineers include:

1. How do scientists obtain the information they need, i.e., what kinds of information resources do they use at different stages of planning and conducting their research and reporting on research results when working independently?

2. How much time is spent in personally using different information and communications resources?

3. How much time is spent in communication with local colleagues?

4. For what kinds of information does the individual investigator depend on assistants, graduate students, etc., and how do they obtain the information?

5. How does information seeking change when working in collaboration with others, and what communications resources are used in carrying out collaborations? With whom does the scientist/engineer (a) communicate and (b) collaborate?

6. To what extent are scientists hampered in addressing research questions because of limitations in access to needed information, and what is the nature of the limitation(s)?

7. With what kinds of electronic and computer-based information/communications technologies and resources are scientists and their students familiar? Which do they know how to use? Which do they prefer?

Some questions that should be addressed to institutions include:

1. What is the role of institutions, departments, centers, etc. in providing information resources?

2. What kinds of electronic and computer-based communication technologies, not including library resources, are available to all scientists/students?

3. Are communication technologies available to which access is restricted to subsets of the population of scientists/students? Which groups?

4. What short- or long-term plans does the institution have for improving availability or access to information resources?

5. What kinds of opportunities are made available to faculty/students for training in the use of electronic and/or computer based resources?

An adequate characterization of the academic institution's information and communication resources requires that library resources also be surveyed:

1. What is the role of the library in planning for academic information services and resources?

2. How much of the library's budget is used for computer-based and other new technological information systems?

3. What kinds of electronic and computer-based systems does the library make available? To whom?

4. How are information systems accessed--by users, by library-based specialists, or both?

5. What kinds of training does the library provide in the use of electronic/computer facilities for library service workers, faculty, researchers, students?

6. To which national databases does the library provide access and at what cost?

7. Is the library formally associated with any other library or libraries; who is included and what is the nature of the association?

SURVEY DESIGN CONSIDERATIONS

The value of current practice information increases as potential sources of variation in behavior are identified and taken into account. Relevant sources of information about the social, psychological, and economic characteristics of the study populations and their institutions, departments, etc., must be reviewed and used in sample selection and data analyses, thus permitting the identification of correlates and possible determinants of the surveyed behaviors. The more precisely such correlates and possible determinants are identified, the more readily subsequent, more detailed investigations of the dynamics that underlie patterns of behavior can be pursued. Relationships between variables revealed by the survey data should help to stimulate and to focus laboratory and experimental studies aimed at achieving understanding of the dynamics of information utilization. It might be possible to generalize beyond specific study populations in subsequent research, even if such research is based on relatively small (but carefully defined) samples. It would be impractical to attempt to study all types of research environments simultaneously because the relevant types vary widely and present different kinds of planning and data collection problems. While industrial involvement in non-defense-related research is increasing, most basic scientific research is still conducted in the academic environment. Furthermore, knowledge about the conduct of academic scientific research is much more abundant and more readily available than is information about industrial research; a great deal more information about the academic environment itself is also known and available. Investigation of the information-seeking behaviors and their correlates among academic scientists may well provide guidance for studies of those behaviors in other environments. For these reasons, it seems prudent to focus initially on research-intensive universities and their faculties.

Within the academic community there exists no formal, federal organization of library services to meet the information needs of scientists in fields other than medicine and agriculture. Other disciplines lack a clearly defined structure and locus of responsibility for service. In the absence of central coordination there is no assurance that any library possesses specific needed information, despite the existence of several networks and consortia that permit the rapid transmission of information once it is identified and located. These circumstances suggest that the study of information-seeking behavior and problems would be sufficiently different for medical and nonmedical scientists as to require quite different study designs and possibly even different approaches. The discussion in this document is addressed to nonmedical sciences.

Institutional and Individual Characteristics in Survey Design

Some of the variables (on which data are readily available) that should be considered in sample selection or for use in analyses of the information user survey include:

Institutional Characteristics[7]

Institution Type (Doctorate, Comprehensive, Liberal Arts, etc.)
Geographical Location
Type of Control (Tax, Private, Combined)
Science/Engineering (S/E) Enrollment over time (o.t.)
S/E Number and level of Degrees Awarded o.t.
S/E Number of Graduate Students (f.t., p.t., postdocs)
Total Expenditures for Scientific Research o.t.
Govt. support through Research Grants, Contracts, and for Training, Fellowships, Construction, Maintenance
Publication/Citation level o.t.
Capital Equipment Expenditures (surveyed in 1982-83)[8]

Characteristics of Departments Within Institutions

Enrollment and Support for Graduate Education (NSF, NIH, Other Govt., Foundations) o.t.

---

[7] Most of the items listed are available on tape or in published form in periodic National Science Foundation reports such as Science Indicators, published biennially, and Academic Science/Engineering: Graduate Enrollment and Support, published annually.

[8] Division of Science Resources Studies, National Science Foundation. Academic Research Equipment in Selected Science/Engineering Fields, 1982-83: An Analysis of Findings from the Baseline National Survey of Academic Research Instruments and Instrumentation Needs. Prepared by Kenneth Burgdorf and Howard Housman, Westat, Inc., 1985.

Doctoral Degrees Awarded o.t.
Faculty Size and Research Support o.t.
Equipment Expenditures, Availability, and Needs (1982-83, see above)
Reputational Ratings of Research Doctoral Programs (1981)[9]

Characteristics of Disciplines

Publication Frequency/Citations in discipline-identified journals
Undergraduate and graduate enrollment and degrees awarded

Characteristics of Individuals

Age
Years since Doctoral Degree
Amount and Source(s) of Research Support
Publication and Citation records o.t.

Variations in Geographic Location, Financial Resources, and Other Characteristics of Institutions

Early (1960s) studies of scientists' sources of information focused on the following resources:

-- Local colleagues and students

-- Non-local colleagues

-- Meeting presentations

-- Preprints/Technical Reports

-- Journals

-- Books

---

[9] Committee on an Assessment of Quality-Related Characteristics of Research-Doctorate Programs in the United States; Lyle V. Jones, Gardner Lindzey, and Porter E. Coggeshall, eds., An Assessment of Research-Doctorate Programs in the United States. Washington, D.C.: National Academy Press, 1982.

The same resources are available today, but the ways in which the written materials and non-local colleagues can be accessed, the speed with which the existence of resources can be discovered and obtained, and the range, kinds, and amounts of information that can be accessed regardless of location or professional status are markedly different. It is, however, highly unlikely that the new technologies such as electronic mail, personal access to bibliographic indexes and computerized files, access to specialized files, networks and bulletin boards, etc., are equally accessible to scientists of different status and in different locations. Even the traditional resources like volumes of journals are becoming unequally available because increased costs are affecting the ability of many libraries to acquire and maintain their collections.

The faculty and students in institutions that are located near technology development centers like Silicon Valley in California, and those in institutions with large budgets for science, may have had greater opportunity in recent years to gain access to the many new technologies and to learn how to use them. Yet little is now known about this possibility. Nor do we know how covariation among intra-institutional characteristics may be related to the ways in which faculty, postdocs, and graduate students are meeting their information needs.

Cross-Disciplinary and Departmental Information

Studies conducted during the 1960s revealed that there were clear differences among scientists in different disciplines in the extent of their dependence on and preference for different types of information resources. There were large differences among disciplines in the rate and intensity with which access and the capability to use computer-based technologies were made accessible during the 1950s and 1960s. Physical scientists in general quickly developed such enormous appetites for computational resources that separate resources had to be developed for various groups, e.g., specialists in the biomedical sciences. It is reasonable to assume that as the many new technologies that have appeared during the last quarter century have become differentially available, patterns of utilization of information resources will have diverged even more.

Some physicists claim they no longer make use of traditional library-based information resources. They see themselves as relying, rather, on (a) direct communication with colleagues through electronic mail, (b) network access to extensive, discipline-specific preprint and reprint resources such as are made available by the Stanford Linear Accelerator Center (SLAC), also through electronic mail, and (c) information resources they develop for their own use. The accuracy and extent of applicability of these claims should be determined. Physical scientists' skills, interests, and research problems have naturally led to their occupying (with computer scientists) a pioneering position in the development, application, and use of computer-based technologies. It can be hypothesized that as pioneers, physicists' information-seeking behavior in 1990 presages patterns of activity that will be seen in other disciplines in future years. Technological advances have had different kinds of effects on the actual conduct of research by scientists in different disciplines. To what extent their needs for and utilization of information resources also differ remains to be seen. It is important, furthermore, to learn how and in what ways the information-seeking behavior of scientists today differs from that of scientists who were active in the 1960s.

The selection of precisely which departmental and disciplinary descriptive variables should be used in sample selection and analysis requires some preliminary investigation. Administrative structures dictate that much of the information that effectively describes differences in research setting is contained in and reported nationally by academic department. Many of the dimensions of size, including research expenditures, availability of instrumentation, number of faculty, postdoctoral researchers, graduate students, etc., are recorded in these terms. Productivity indicators, on the other hand, are most often measured in terms of discipline. Also, while departments and disciplines are to a large extent coextensive, many individual scientists today are employed in departments whose primary discipline is different from their own. How these "anomalies" may affect either the individual or others in the department is not known, but the circumstance is one that should be looked into.

Variations in Age and Productivity

Studies of relationships between age and the productivity of scientists have indicated that relations between these variables are not simple, either within or between disciplines. It is unlikely that patterns of information-seeking behavior vary directly with age either, but this question should be explored, also drawing into consideration the scientist's attitudes toward electronic information and communication media. Elucidation of the extent to which patterns of information-seeking behavior are related to age, attitudes, and productivity (e.g., as measured by publications, citations, and patents) would provide potentially important indications of (a) how the uses of resources are likely to change as the information technology backgrounds of younger investigators changes, (b) whether productivity affects or is affected by the availability and uses of information resources and technologies, and (c) the possible importance of issues of availability, access, and attitudes toward resources.

For large-scale studies, publication data provide the least biased and most widely available source of information about the productivity of institutions, disciplines, and individuals. In contrast with other variables under discussion, however, the availability of data is controlled by commercial interests. Cost considerations may therefore complicate decision making about the nature and extent to which needed information can be obtained and used.

PRELIMINARY STUDIES AND PLANNING

Before the design of a wide-ranging academic survey can be completed, some preliminary steps and investigations are needed:

* A search of recent literature. In 1982 the National Science Foundation published an annotated bibliography in which studies of scientific disciplines published before mid-1980 were listed.[10] Nearly fifty studies involving "information exchange" are listed. This material and the isolated studies conducted since 1980 that address relevant information-seeking and information-exchange behaviors should be reviewed for the possible identification of useful sampling or analytical variables.

* A series of exploratory interviews with a variety of physical scientists. The range should be broad because little is known about how different disciplines and subdisciplines use information resources. For example, it has been alleged that some physicists and mathematicians make little use of traditional resources, while chemists are said to use online abstracting services extensively. The sample should be sufficiently large to reveal whether markedly different patterns exist, and whether

methods of data gathering other than mail questionnaires will be needed to ensure that useful information will be obtained from a national sample.

---

[10] Division of Planning and Policy Analysis, Office of Planning and Resources Management. Studies of Scientific Disciplines: An Annotated Bibliography. Washington, D.C.: National Science Foundation, 1982.

* Appointment of a survey design advisory group. In order to maximize the value and utility of a broad survey of the information-seeking behavior of scientists from several academic disciplines, it would be advisable in the process of completing a study design to obtain the assistance and guidance of experts in at least four areas: (a) experts in sample selection, questionnaire design, and data analysis of large scale surveys; (b) experts in the design and analysis of studies of the social science of science issues who have had experience in working with the scientists of the several academic science disciplines that are of interest; (c) experts in developing and providing library and other scientific information resources and services; (d) experts in the development and use of computer and communications technologies. The participation in such a group of a representative of each of the broad groups of disciplines to be involved in the study also would be advisable.

CONCLUSION

The proposed survey should produce information that would be useful to all federal research support agencies, to national scientific professional and research organizations, to higher education institution department chairs, administrators, and librarians, to social scientists whose research is concerned with achieving understanding of the dynamics of scientific effort and achievement, to commercial organizations dealing with information resources, to the current and future scientists in whose interest the information is sought, and to the National Science Foundation (NSF) and its Science Indicators Unit. To wit:

-- Federal research support agencies such as the Department of Agriculture (DOA), Department of Defense (DOD), Department of Energy (DOE), Health and Human Services (HHS), and NSF will acquire knowledge about current circumstances surrounding the use and availability of information resources that will permit better informed decision making concerning policies and procedures related to current and future funding for such resources.

-- National scientific professional and research organizations will have access to state-of-the-science data concerning their relevant disciplines, which should assist in assessing needs for program development. They also will have a basis on which to consider current and potential needs for the development of new approaches to meeting the information and communication needs of their constituent scientists.

-- Higher education institutions will learn how information resource availability, access, and services in their institutions compare with those of others. The information can be used to assist in assessing needs--e.g., for training and for the improvement or acquisition of needed resources and services.

-- Social scientists will have a knowledge base that should aid the efficient planning of well-focused studies of the dynamics of interrelationships between information resources, services, communications, performance, and productivity.

-- Information industry firms will have available a snapshot of the current academic market, and hints of potential markets.

-- Individual scientists will learn how their own patterns of behavior compare with those of others in their own and other disciplines, and will be able to compare their opportunities for access to information resources with those of their colleagues in the same and other types of institutions.

-- The National Science Foundation and its Science Indicators unit will have an initial set of estimates of a new group of measures that may serve as a useful new component in describing the status of the American scientific enterprise. The survey would, of course, mark only the beginning of what should become an enlarged NSF program of supported psychological, sociological, and communications research focused on achieving deeper understanding of how information processing affects and is affected by the characteristics of individuals and environments. Coupled with subsequent research that is likely to reveal more than is now known about the dynamic interplay between environments, behavior patterns, and productivity, Science Indicators could acquire a broader dimension that would be useful to the entire scientific community. Thirdly, periodic follow-up research on the more significant sources of variation in information-seeking behavior and resource availability and access would keep NSF abreast of ongoing changes in information patterns.

The importance of learning more about the users and uses of information technology was recognized in a recently released report of the NAS Committee on Science and Public Policy (COSEPUP) in which "the users' view" of information technology and the conduct of research was discussed. The report states,

> ...there is almost no systematic information on the users and uses of information technology.... [We] cannot estimate how many or what proportion of scientists use computers in different fields, how access to networks and computer facilities is distributed across disciplines, or to what extent useful applications are disseminated throughout the research community. Systematic collection of such information is essential to the development of intelligent policy. Researchers' experiences in using information technology can help guide decisions about policy and resource allocation. In turn, these decisions will shape the technological and institutional advances that break down impediments to the further use of information technology. This process will continue to change the nature of scientific, engineering, and clinical research itself.[11]

It has been said that easy access to computer networks and fax technology has facilitated the transfer of knowledge to the point where journals and even conferences are obsolete as the primary modes of information transfer. While this may be partially true in reference to the immediate transfer of recently developed information among certain subgroups of scientists, the idea must be challenged in reference to (a) the circumstances that surround the work of the majority of active academic scientists, and (b) the exchange of information that is more than one or two years old. The fact is that, apart from a few special studies, we do not know how the majority of scientists today are meeting their information needs, and until the facts are known it is not expedient to make major plans for "improvement."

---

[11] National Academy of Sciences, Report of the Panel on Information Technology and the Conduct of Research, Committee on Science, Engineering, and Public Policy (Donald N. Langenberg, Chair). Information Technology and the Conduct of Research: The User's View. Washington, D.C.: National Academy Press, 1989.

This document reflects a conviction that increased knowledge and understanding of how information resources are actually being used by scientists and engineers, and how utilization of scientific information is affected by advancing technology, are matters of paramount importance to the users, their institutions, and policy makers at national levels. The discussion outlines some of the considerations that need attention in planning and carrying out the first steps of a full inquiry.

Library Resources and Research Productivity in Science and Engineering: Report of a Pilot Study

Prepared by

Nancy A. Van House
School of Library and Information Studies
University of California
Berkeley, California

Prepared for

Council on Library Resources
Washington, D.C.

July 24, 1990

ABSTRACT

This was a pilot study testing for correlations between scientific and engineering library resources and scholarly productivity in a small judgmental sample of research universities. Science and engineering were defined as "the biological and physical sciences, mathematics, engineering, and computer science, not the social or behavioral sciences." Medicine was excluded. Data were collected on science library resources, and on research productivity measured by publications and faculty honors.

The major finding was that library resources in science and engineering are correlated with faculty productivity. More science serials, larger science collections, and more professional science librarians correlated well with more publications per faculty member.

Are these correlations simply a function of university characteristics that correlate with library resources? After data in the study were controlled for size (measured by total faculty) and research funding, library resources still affected research productivity.

The library resource measures were themselves intercorrelated, making it difficult to sort out their relative influences. The greatest surprise, however, may be the large correlation of research productivity with professional library staff. Further investigation is needed to determine whether this is simply due to the relationship between staffing and other resources. Many academic library users underestimate the role of professional library staff in building and maintaining the collection and in helping clients make maximum use of the collection and locate information elsewhere.

Another important finding was the decline in purchasing power among the sample libraries when expenditures are corrected for increases in materials prices.

The main implication of this study is that science library resources may well play a role in science research productivity. If so, the possible effects of declining library purchasing power could be serious. Although this study cannot prove a causal connection, it does indicate a need for concern. If library resources do affect research productivity, it may soon be too late to undo the damage from missed acquisitions and deteriorating collections.

The author would like to acknowledge the assistance of Beth Weil, Annette Melville, David Sullivan, Diana Laslett, and Dan Dabney.

INTRODUCTION

This is the report of a pilot study addressing the question: do library resources contribute to research productivity in science and engineering? Data on science library resources were collected from a small sample of universities and correlated with data on their research productivity.

The study was conducted on a limited budget and tight time frame with a small sample. Data were limited to those readily available from the sample institutions and other relevant sources. This was truly a pilot study designed to determine whether there is sufficient evidence for a larger study, to identify the major issues to be addressed in such a study, and to help in its design.

The results must be considered in light of the study's limitations. However, the findings suggest that this is a fruitful avenue of research to pursue. There is indeed evidence that library resources correlate with faculty research productivity.

THE DETERMINANTS OF RESEARCH PRODUCTIVITY

A major question in science policy is whether greater resources increase the rate of science knowledge growth. This can be asked at the national level: does national R&D spending result in knowledge growth? It is also asked at the institutional level: What are the characteristics of institutions that produce good research? What is the relationship between institutional resources and research productivity?

One approach to research productivity and knowledge growth (Cohn, 1986; Holzer, Dunn, and Shahidullah, 1987) is the input-output model, which describes the marginal effects of inputs on research outputs using a production function. This research seeks to identify the appropriate measures of knowledge growth and the significant inputs, and describe the relationships among them. The exact form of the production function is likely to vary across fields.

Cohn (1986) hypothesizes that the rate of knowledge growth in the sciences depends on the following factors:

1. The adequacy of the existing knowledge base
2. The number of scientists with skills and ability to locate and solve important problems given the current knowledge base
3. The availability to these scientists of resources, including time, instrumentation, technical assistance, and

organizational structures for communication and coordination.

Empirical tests of the links between resources and scientific productivity have had mixed but encouraging results (Cozzens, 1986a). We have a long way to go in understanding the appropriate measures of research outputs, the relevant inputs, and the relationships among them.

LIBRARIES AND RESEARCH PRODUCTIVITY

The library is an important resource for research. Libraries provide access to the knowledge base and play an important role in communication and coordination among researchers (in Cohn's terms). They provide both physical and intellectual access to the published literature: physical access through copies of publications, and intellectual access through finding tools, which inform the researcher about potentially relevant publications.

Publications serve at least three major functions in the sciences: to record observations and findings, to enable others to duplicate and validate research, and to provide evidence of scholarly achievement. The first two functions are part of the communication system. The third is part of the reward system. Publications are used as evidence of scholarly achievement in, for example, the academic tenure and promotion system.

The uses of published literature for communication vary across fields. In some, such as physics, the research front moves rapidly and scientists rely on preprints. By the time an article appears in print it is already old news. In other fields, however, researchers rely more on published literature. In mathematics, for example, some of the most significant problems are those that have remained unsolved for a long time, and so the literature concerning them has a long life.

Dependence on the published literature also varies with individuals. Scientists often rely on the "invisible college" (colleagues, informal contacts, preprints, etc.) for news of relevant research.

The library as a source of access to the published literature is probably most important to the following groups:

1. Students, especially graduate students (who often account for the majority of academic library users)
2. Beginning scientists
3. People working outside their immediate fields (and much creative work consists of making links across disciplines and specialties)

4. Those who are more isolated from the invisible college (due to location at a less prestigious institution, lack of funds for travel, etc.)

The library is not the only source of the published literature. People subscribe to journals and buy monographs; authors distribute preprints and reprints; colleagues pass along photocopies. Given the volume of materials published and the high prices of scientific journals, however, a researcher can acquire only a fraction of the important publications in his or her immediate field. He or she still relies on the library for other materials. Students and researchers with limited funding are at a disadvantage in buying their own copies.

In recent years, developments in technology and in resource sharing have phenomenally expanded the research library's scope. A library's own collection is only the beginning. The growing variety and accessibility of bibliographic databases make it easy for the researcher to discover what's been published. Even the smallest libraries are now able to acquire virtually any publication on demand through interlibrary loan and document delivery services.

Yet virtually no empirical evidence on the library's contribution to research exists. The only study to empirically investigate the effect of library resources on research programs was the National Academy of Sciences study of the quality of science graduate programs (Jones, Lindzey, and Coggeshall, 1982). They correlated objective and subjective measures of program quality, program descriptors, and resources by discipline. One measure was a university library size index derived from a factor analysis of data from the Association of Research Libraries statistical reports.

Among six science disciplines, library size exhibited a moderately high correlation (.6 to .7 among the different disciplines) with reputational measures of the scholarly quality of faculty. It exhibited a similar correlation (.4 to .7) with the number of publications by faculty. Its correlations with university research expenditures in a discipline, however, were fairly low: except for one .45, all were .33 or below.

THIS STUDY

The present study tested for correlations between scientific and engineering library resources and scholarly productivity in a small sample of research universities. Science and engineering were defined as "the biological and physical sciences, mathematics, engineering, and computer science, not the social or behavioral sciences." Medicine was excluded. Data were collected on library resources and on research productivity as measured by publications and faculty honors. The sample is a small judgmental sample of major university libraries.

The level of analysis is institutional. Most studies of research productivity have been at the subdiscipline level (Cozzens, 1986b). Production functions probably differ by discipline or subdiscipline (Cohn, 1986). Subjective rankings of research and/or educational quality are most appropriately done at the discipline or departmental level.

Empirically, however, it is difficult to find equivalent measures of research productivity and library resources at the discipline and subdiscipline level. In the past, few libraries have had data on library resources by subject classification (Machlup, 1978-80). This is changing, but such data as do exist are often not comparable across libraries, and the statistical categories rarely match academic departmental boundaries.

As libraries computerize their catalogs and acquisitions records, analyses by subject area are increasingly feasible. The Research Libraries' Group is sponsoring the RLG Conspectus, library self-ratings of collection intensity by subject area. These are self-assessments, however, so comparability across libraries is questionable. The National Shelflist Count is another attempt to collect data on research library resources by subject area, but currently only a handful of libraries are represented.

Given the problems of matching data, the only feasible level of analysis for the present study was institutional, though future studies at the level of department or discipline would be worthwhile.

THE SAMPLE

The sample was a judgmental sample of U.S. universities whose libraries belong to the Association of Research Libraries (ARL).[1] ARL membership is by invitation only. Members are North America's largest research libraries. Of the 119 members, 107 are academic libraries.

ARL member institutions were chosen for two major reasons. First, our primary interest in this study is research universities. Second, using ARL libraries simplified data collection. We based our library data collection on ARL's annual survey of its members' library resources. Given the scope and time frame of this study, it was important to have comparable data readily available for the sample libraries.

The universities included were chosen based on the following criteria:

- They were distributed throughout the ARL rankings, from the largest to the smallest libraries.
- An attempt was made to include both institutions with major science research reputations, and those without.
- They were distributed through the rankings of universities receiving NSF research funds.
- Some were included because we had reason to believe that their libraries would have the data that we requested.

LIBRARY DATA

Data on library resources in science and engineering were collected by a survey. A questionnaire was designed based on the annual ARL statistical survey (Association of Research Libraries, 1990). The ARL data elements were used as the starting point for three reasons:

- The definitions and measurement methods appropriate to the major academic research libraries have already been debated and decided upon by ARL.
- If needed, comparable library-level data would be available for the sample libraries from the annual ARL statistical reports.

---

[1] Two non-ARL libraries were added to the sample because of their universities' stature as centers for science research.

- o Since ARL members collect library-level data using ARL definitions, if they have data on science and engineering resources they are likely to use the same data elements.

The questionnaire and cover letter are in Appendix A. The line numbers of our questionnaire corresponded to the ARL survey to show respondents the parallels between the two data collection forms. ARL definitions of the data elements were included with the questionnaire.

Several library resource categories were added to the survey in the hope of broadening the scope of resources considered. They are labelled with letters instead of arabic numbers on the questionnaire. These represent new types of information resources that are changing the means by which scientific and technical information is accessed and stored (bibliographic databases are one example). However, the responses in these added categories were insufficient to include them in the final analyses. This is unfortunate, since these are important and growing library resources. Future studies of science and engineering library resources must include these new areas. However, unless ARL defines these data elements, research libraries will not collect consistent and usable data.

The data elements included in the final analyses were:

Science/engineering volumes held (including monographs and bound volumes) as of June 30
Science/engineering volumes added during the fiscal year, net (additions minus volumes retired or declared lost)
Current science/engineering serial titles received, purchased and not purchased (some titles are received on exchange)
Professional science/engineering library staff, FTE
Annual science/engineering monograph expenditures
Annual science/engineering serial expenditures
Science/engineering faculty

One data element that might have been included but was not due to the lack of comparable data across libraries was use, as measured primarily by circulation. Methods of counting circulation vary, and, since it is not included in the ARL survey, responding libraries were unlikely to be able to provide consistent data.

Data were collected for academic years 1984-5 through 1988-9. Not all libraries could provide all the requested data, so for individual items the number of responses varies.

The questionnaire was drafted, pretested on several librarians at UC-Berkeley and Stanford, and revised. Questionnaires were sent by telefacsimile to 43 sample libraries on March 28, 1990. They were sent to the heads of science libraries, if possible; otherwise they were sent to heads of collection development, or, if all else failed, to the director of the library. A response was requested by April 13, 1990. Libraries were given the option of sending their internal reporting forms instead of using the questionnaire.

During the week of April 16, nonrespondents were telephoned by the investigator.

The final response was as follows:

	Number	Percent
Questionnaires returned and usable:	27	63%
Returned, not usable:	1	2
Declined to participate/no answer:	15	35
Questionnaires distributed:	43	100

The names of participating institutions appear in Table 1.

Sixty-three percent is an excellent response. Many who declined expressed interest but lacked data on science and engineering resources separate from total library resources. Several libraries did considerable processing of internal data to generate the figures requested. Some respondents noted that the quality of data was not as good as it might have been if it had been collected prospectively. Participants said that although their data were reasonably accurate, their figures were often estimates, and asked that library-level survey data not be reported.

Table 1
Research Productivity Rates by Institution
1989

Institution	per Science Faculty Member Sci Citation Index Publications	Awards NAS+NAE+PYIs
CAL TECH	5.95	.45
COLUMBIA	5.82	.20
STANFORD	3.96	.43
UC BERKELEY	3.69	.25
MIT	3.47	.31
INDIANA	2.84	.04
UNIV OF UTAH	2.61	.05
PURDUE	2.58	.06
SUNY BUFFALO	2.55	.03
NYU	2.54	.09
DUKE	2.47	.06
PRINCETON	2.36	.18
PITTSBURGH	2.33	.02
UNIV OF MICHIGAN	2.32	.06
UC DAVIS	2.05	.02
UNIV OF WISCONSIN MADISON	1.86	.06
BROWN	1.65	.04
YALE	1.49	.10
RICE	1.37	.04
WAYNE STATE	1.34	.00
UNIV OF FLORIDA	1.31	.01
SO CAROLINA	1.11	.00
ARIZONA STATE	1.02	.02
GEORGIA TECH	.83	.03
SOUTHERN ILLINOIS	.69	.00
RUTGERS	*	*

Spearman's rank order correlation between publications per faculty member and awards per faculty member: .76

*Number of science faculty not reported.

RESEARCH PRODUCTIVITY

A number of studies have ranked universities and programs according to the quality of their research or of their graduate education, including Cartter (1966), Roose and Andersen (1970), and Jones, Lindzey, and Coggeshall (1982).

Two types of measures have been used, subjective and objective. Reputational studies reply on subjective, peer ratings of individual programs. This method was used by the widely-cited Cartter (1966) and Roose-Andersen (1970) studies. Its major advantage is that it is based on the expert judgment of knowledgeable scholars. Its major disadvantages are that subjective assessments may be based on old or incomplete information, and evaluators' criteria may be varied and uncontrollable.

Objective measures eliminate the subjectivity of reputational studies, but are limited to criteria for which measures can be defined and reliable data collected across institutions or programs. A wide range of measures are possible, each reflecting a different aspect of the program. The choice of measures is critical to the outcome of the study. The Conference Board study (Jones, Lindzey, and Coggeshall, 1982) used 16 measures of program size, characteristics of graduates, university library size, research support, and publication records, as well as reputational surveys, to assess research doctorate-granting programs in the physical and mathematical sciences.

Most assessment studies have been done at the level of discipline. These have the advantage of addressing issues specific to each discipline. For reputational studies, program or discipline is probably the level at which experts can make the most valid judgments.

For this study, the level of assessment was the university, specifically, university science and engineering programs. Institutional research productivity was measured using two kinds of objective measures: researchers recognized as outstanding, and publications.

Outstanding Researchers

Three measures of outstanding researchers were used:

1. Number of National Academy of Sciences (NAS) members as of January 1990. NAS members are elected by other members based on outstanding achievements in science. A list names and professional affiliations of its members was obtained from NAS.

2. Number of National Academy of Engineering (NAE) members as of February, 1990. NAE members are elected based on having made important contributions to engineering theory and practice, and/or having demonstrated unusual accomplishments in the pioneering of new and developing fields of technology. A list of names and professional affiliations of its members was obtained from NAE.

NAS and NAE membership recognizes outstanding scientists and engineers. However, NAS and NAE membership comes after a researcher is established, often well into his or her career. Election may be a better indicator of past than present achievements. To compensate for this bias and include younger and currently active outstanding researchers, a third measure was added.

3. Number of NSF Presidential Young Investigators (PYIs). The PYI awards recognize outstanding young scientists and engineers at the beginnings of their careers. They are selected on the basis of demonstrated ability and potential for contribution to science and engineering effort. An annual summary of PYIs by institution was obtained from NSF. This listing did not distinguish them by field of study, so some researchers in the behavioral sciences were included.

Publications

Publications were measured by articles indexed in Science Citation Index (SCI). SCI is the single most comprehensive index to the literature of the sciences and engineering.

SCI was searched by institutional affiliation of author(s) to determine the number of articles with at least one author from each of the sample institutions for the years 1985 to 1989. Articles by authors from more than one institution were counted once for each institution. Articles by more than one author from the same institution were counted once for that institution. Included were articles, literature reviews, and bibliographies. Excluded were notes, book reviews, and abstracts. Also excluded were articles that could be identified as authored by medical, dental, or veterinary school faculty.

Other researchers have done more to manipulate SCI publication data; for example, counting an article with multiple authors as a fraction of a publication for each author, or differentiating between more and less influential publications. Aside from the usual conceptual and methodological questions about such rearrangement of the data, practical considerations precluded making such adjustments for the thousands of articles included in this study.

FINDINGS

Because this was a judgmental and not a random sample, representing a significant proportion of ARL membership, the results apply only to the responding libraries. Tests of significance infer the parameters of a larger population from sample statistics. We are not inferring results for a larger population but simply reporting results for the responding libraries, so tests of significance are not appropriate. However, respondents were widely distributed among ARL libraries in resources and research productivity, suggesting that these results might hold for the universe of ARL libraries.

Library Resources

Table 2 reports the means and standard deviations on the science library resource variables for responding libraries. The measures of science library resources were highly correlated. Table 3 presents the 1989 correlations; other years were similar. Holdings, serial subscriptions, professional library staff FTE, and bound volumes added correlated with one another in the range of .72 to .84, indicating that a library that is large on any of these variables tends to be large on all of them.

Correlations with the expenditure variables were lower. We found considerable annual fluctuation in expenditures within libraries. University budget crunches often affect the library budget, particularly the amount of money available for monograph expenditures. Other resource variables respond more slowly to budget fluctuations.

The generally high correlations in this study mean that we can use a single measure or a small number of measures to stand for library resources. Although it would be preferable to sort out the effects of different resources on productivity simultaneously, our sample is too small to do so.

A notable finding of this study is that responding libraries' purchasing power has been seriously eroded during the five years considered (Figure 1). This verifies the findings of other studies (for example, Association of Research Libraries, 1990). The rapidly rising cost of library materials, especially journals, is of major concern to research libraries. A number of recent studies have addressed the causes and the impact on libraries of this marked increase in prices (for example, Association of Research Libraries, 1989).

Correcting for changes in the price of serials (see Appendix B for method), average serials expenditures among the libraries in our sample actually declined 7 percent from 1984-5 to 1986-7, after which they increased but remained below the 1984-5 level (Table 4).

Table 2
Science Library Resource Measures by Year

Mean
Standard Deviation
No. of Observations

	1985	1986	1987	1988	1989
Net Additions	16430	13810	15880	14530	14440
	10050	10180	13150	13270	7290
	(14)	(18)	(18)	(18)	(20)
Holdings (bound vols)	475500	484680	479410	509900	532650
	283000	301340	301340	317060	308050
	(16)	(17)	(18)	(18)	(22)
Serial Subscriptions	6970	6910	6870	6880	5900
	6350	6230	6340	5530	3930
	(13)	(14)	(14)	(14)	(22)
Professional Staff FTE	7.5	7.3	7.4	9.6	9.6
	4.9	4.4	4.5	8.5	8.0
	(20)	(21)	(21)	(23)	(26)

Table 3
Correlations among Library Resource Measures
1989
(no. of observations)

	Net Additions	Holdings (bound vols)	Serial Subscriptions	Prof'l Staff FTE	Monograph Expenditures
Holdings	.77 (20)				
Serial Subscriptions	.72 (19)	.80 (21)			
Prof'l Staff	.73 (20)	.83 (22)	.84 (22)		
Monograph Expenditures	.54 (20)	.44 (22)	.35 (22)	.48 (26)	
Serial Expenditures	.55 (18)	.43 (20)	.64 (21)	.49 (24)	.73 (25)

Figure 1

Mean Library Materials Expenditures
1984-85 to 1988-89, in 1985 dollars

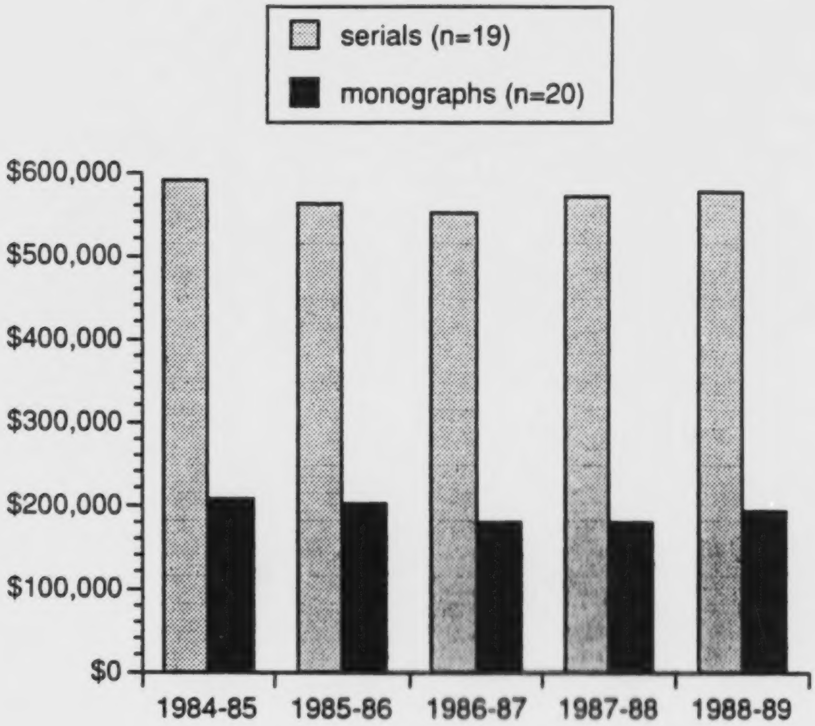

Table 4
Average Serials Expenditures, 1984-5 to 1988-9
(n = 19)

Year	Current $	1985 $
1984-5	$590,700	$590,700
1985-6	647,800	563,800
1986-7	768,800	551,500
1987-8	869,000	571,000
1988-9	958,500	576,700

Table 5
Average Monograph Expenditures, 1984-5 to 1988-9
(n = 20)

Year	Current $	1985 $
1984-5	$208,000	$208,000
1985-6	219,200	201,700
1986-7	219,600	180,900
1987-8	235,100	179,900
1988-9	272,000 (est.)	193,300

For monographs, as well, real expenditures declined, then increased. In 1987-88 they were still 14 percent below 1984-5 levels (Table 5).[2]

Research Productivity

The measures of research productivity were: members of the National Academies of Sciences (NAS) and Engineering (NAE), NSF Presidential Young Investigators (PYIs), and science publications. Table 6 reports means and standard deviations for the institutions in our sample. The variable "awards" refers to the sum of 1989 NAS and NAE members plus PYIs for the five years.

The absolute values of the productivity measures were moderately to highly intercorrelated (Table 7). Publications are extremely highly correlated across years, consistently at about .99: the universities that publish the most in the sources indexed by SCI do so consistently from year to year.

The remarkable stability of publications over time may be due to several factors. First, articles published in one year were written at different times, so factors affecting publication at various times and in various disciplines may balance out. Second, faculty productivity may be relatively insensitive to short-term environmental factors. A decline in university research support, for example, may only slowly affect publication. In the short run, faculty may be inconvenienced but compensate. In the long run the impact may be greater. Distinguished faculty may move to more supportive institutions; fewer research projects may be initiated; graduate students may go elsewhere.

The various types of productivity measures were moderately to highly correlated. The universities that produce large numbers of publications also have large numbers of outstanding faculty (as measured by NAS and NAE membership and PYIs).[3]

Of course, larger faculties produce more publications and win more awards. Table 1 shows the ratios of science publications and of awards to science faculty. Unfortunately, many respondents could provide the number of science faculty only for 1989, so in the analyses that follow publications and awards per science faculty member are only for 1989.

---

[2] Data collected on monograph additions were net, not gross, so we cannot simply use volumes added to measure libraries' purchasing power.

[3] NSF made about half the number of PYI awards in 1986 as in other years, making this year anomalous.

Table 6
Productivity Measures
(n=27)

Variable	Mean	Standard Deviation
Natl Academy of Sciences Members	19.3	25.9
Natl Academy of Engineering Members	11.3	20.8
NSF Presidential Young Investigators:		
1985	3.3	3.9
1986	1.7	2.1
1987	3.1	3.4
1988	2.7	3.1
1989	2.7	3.1
Awards: NAS + NAE + PYIs 1985 through 1989	44.0	57.6
Science Citation Index Publications:		
1985	895	633
1986	909	620
1987	949	625
1988	979	650
1989	1051	690

	Presidential Young Investigators					Academies		Sum		Publications		
	PY15	PY16	PY17	PY18	PY19	NAE	NAS	AWARDS	SC15	SC16	SC17	SC18
PY16	.53											
PY17	.88	.69										
PY18	.69	.71	.75									
PY19	.90	.56	.82	.68								
NAE	.76	.69	.71	.69	.78							
NAS	.78	.77	.81	.77	.79	.88						
AWARDS	.84	.77	.84	.79	.85	.95	.97					
SC15	.80	.54	.75	.64	.84	.69	.76	.79				
SC16	.78	.52	.74	.63	.83	.67	.74	.77	.99			
SC17	.78	.51	.73	.63	.82	.67	.74	.76	.99	.99		
SC18	.79	.52	.73	.63	.83	.68	.74	.77	.99	.99	.99	
SC19	.77	.51	.71	.63	.81	.67	.73	.76	.99	.99	.99	.99

Table 7. Correlations among Research Productivity Measures

Table 8
Correlations among Research Productivity Ratios: Productivity Measures Divided by Science Faculty

SCI Pubs per Science Faculty	Awards* per Science Faculty	Science Citation Index Publications per Science Faculty			
		1985	1986	1987	1988
1985	.74 (7)				
1986	.69 (7)	.99 (27)			
1987	.70 (8)	.99 (27)	.99 (27)		
1988	.69 (8)	.97 (7)	.96 (7)	.99 (8)	
1989	.79 (26)	.98 (7)	.98 (7)	.99 (8)	.99 (8)

*Members of NAS + NAE + NSF Presidential Young Investigators 1985 through 1989

Correcting for faculty size (Table 8), the productivity measures remain highly correlated. The ratio of science publications to science faculty members remains consistent across the years, although the sample before 1989 is small. The rank order correlation between publications and awards per faculty member is .76, indicating reasonable consistency across the two measures. These ratios must be interpreted with caution. Numbers of science faculty were provided by librarians or campus administrators. The reliability of these data is unknown. An added problem is presented by non-teaching researchers who are not counted in faculty totals. The only way to eliminate their publications from the Science Citation Index sample would be to check laboriously each of thousands of authors against college catalogs and directories, an impossible task given the time constraints of this study.

The high intercorrelations indicate that the productivity measures all represent basically the same thing and that we can use any one of them. As with library resources, we would prefer to use several productivity measures simultaneously, but our small sample precludes it.

The implication of stable publication rates and ratios over time for the present analysis is that we can't expect to find changes in publication rates as a function of library resources within universities over the short time period examined. Instead, we will be looking for differences across universities as a function of resources.

Relationship between Resources and Productivity

The primary question addressed by this study is: is science and engineering research productivity related to science and engineering library resources? Given the stability of the productivity data, we cannot look for time lags to indicate causality. What we can look for is correlations, which do not prove causality, of course, but which may indicate a relationship that merits further investigation.

Appendix C shows the correlations between research productivity and science library resource variables by year. The publications in any given year may draw on library resources of earlier years, so for each publication year earlier years' library resources are included. Scientists do not necessarily use science library resources exclusively, so correlations between research productivity and total library resources were also examined (using data from the annual ARL statistical reports), but the correlations were lower than for science library resources and are not reproduced here.

Overall, total science publications tend to correlate best with science library professional staff FTE, number of current science serial subscriptions, and science volumes held. The correlations between most science library resource measures and publications are fairly stable across publication year, probably due to the stability of both sets of measures.

Additions are less stable. Volumes added were net, not gross, so these figures reflect not just acquisitions but discarded materials. The attention paid by librarians to weeding outdated materials and deleting records for lost items varies, creating fluctuations in net additions.

Given the fluctuations observed in expenditures, especially for monographs, it is not surprising that these correlations tend to be lower. Productivity measures for 1989 correlate more highly with earlier than with later expenditures measured in constant dollars. This probably represents declining real expenditures.

All things being equal, a larger faculty will produce more total publications. To correct for institutional size, science publications were divided by the number of science faculty members. The only year for which we have sufficient data on science faculty is 1989. Since items published in 1989 are based on research done earlier, it is appropriate to correlate 1989 publications with library resources over the five years for which we have data.

It would not be appropriate also to divide science library resources by faculty size. Library resources are not distributed among the faculty as are, for example, office supplies. An increase in library resources usually means greater depth and variety of publications. A single faculty member will have access to a much greater number and range of items in a larger library regardless of the number of other people using the same resources.

From Table 1, it is clear that Cal Tech and Columbia have exceptionally high publication rates per faculty member. Graphing publications per faculty member against library resources also reveals Georgia Tech as an outlier. Correlational analysis requires a homogeneous sample. It assumes a linear relationship and is highly sensitive to outliers, so the usual procedure is to drop outliers. Cal Tech, Columbia, and Georgia Tech were dropped from Table 9 and from subsequent analyses based on science publications per science faculty member. Cal Tech and Georgia Tech are clearly different kinds of institutions than the more broadly based universities in the sample. There are several possible explanations for Columbia's position as an outlier: it may also be a different kind of institution, the reported number of science faculty at Columbia may be erroneous, or it may be anomalous in some other way.

Table 9
Correlations between 1989 Science Publications per Faculty Member and Library Resources 1985-1989

(no. of observations)

	Library Resource Years				
	1985	1986	1987	1988	1989
Net monograph additions	.64 (11)	.33 (14)	.51 (14)	.31 (14)	.16 (16)
Total monograph holdings	.47 (13)	.49 (13)	.44 (14)	.35 (14)	.31 (18)
Serial subscriptions	.56 (11)	.56 (11)	.56 (11)	.59 (11)	.49 (18)
Professional library staff FTE	.51 (19)	.51 (19)	.53 (19)	.39 (20)	.34 (22)
Monograph expenditures 1985 $	.37 (18)	.42 (19)	.54 (19)	.46 (20)	.35 (23)
Serial Expenditures 1985 $	.37 (17)	.43 (17)	.41 (18)	.40 (19)	.43 (21)

Outliers removed.

For the remaining universities, publications per science faculty member correlate moderately with science library holdings, additions, serial subscriptions, and professional library staff FTE (Table 9). Correlations are generally higher with earlier years' resources. This is reasonable considering publication time lags, and tends to reinforce the idea that library resources contribute to publications. (The sample size also varies over time, as more libraries were able to provide more recent data.)

The more science serials, volumes, and professional staff a library has, the more publications are produced per faculty member. Because of our small sample and the high correlations among the library resource variables (Table 3), we cannot disentangle the effects of various library resources. We can say, however, that science library resources and publications per faculty member are correlated.

Are these correlations simply a function of university characteristics that, in turn, correlate with library resources? Are the library resource variables proxies for something else? Two university characteristics that are likely determinants of faculty productivity are size and research expenditures.

To control for the overall size of the institution, publications per science faculty member in 1989 were regressed on library resources and faculty size. The library resource variables were science serial subscriptions, science library staff, holdings of bound science volumes, and annual volumes added. Publications were regressed on each resource variable, one at a time; the small sample precluded entering all the resources variables into a single regression equation. For each resource variable, two regressions were run: one using 1989 data, and the other 1987, which had the highest overall correlations with publications per capita. Also entered into each regression was size of institution as measured by total faculty (not just those in the sciences) in 1989.

In no case was the coefficient on faculty size significant at the .05 level. Because of missing data, the samples for these regressions were small (11 to 18 cases), so the results should be seen as indicative, not definitive. But the evidence from this small sample is that the effect of library resources is independent of that of overall university size.

To control for research expenditures, publications and awards per faculty member were each correlated with three measures of research funding: NSF funding and total federal academic science and engineering support (both for fiscal 1988, the most recent available, from the National Science Foundation, 1990); and total R&D expenditures, FY1987 (National Science Board, 1989). Funding tested was both total and per science faculty member (Table 10). The correlations are generally moderate, but high between

Table 10
Correlations between R&D Resources and Productivity Measures

(n = 23)

	NSF support	Federal academic science/ eng. support	Total R&D expenditures
Correlations with Science Publications per Faculty Member, 1989:			
Total $	.66	.40	.26
$ per science faculty	.90	.58	.62
Correlations with Awards* per Faculty Member, 1989:			
Total $	.77	.57	.43
$ per science faculty	.83	.55	.60

*Awards = NAS members + NAE members + PYIs 1985 through 1989 publications and awards per capita and NSF support per science faculty member (.90 and .83).

Table 11 relates research expenditures to library resources.[4] Total funding correlates moderately, at best, with science library resources; funding per science faculty member is basically unrelated to science library resources. This suggests that the influence of library and research expenditure variables on research productivity would not be redundant.

To control for the effect of funding on productivity, a series of regressions were run relating publications and awards per faculty member to library resources and research funding. For the library resource measures, data were for 1987, the year that correlates best overall with 1989 publications. Again, with missing data the samples were rather small for regression analysis, so the results should be considered indicative, not definitive.

Tables 12 and 13 summarize the results. Publications, funding, and library resources together have quite good explanatory power, with R^2's (that is, proportions of variance explained) in the range of .36 to .66. Even when funding is taken into account, holdings, serials subscriptions, and especially professional library staff remain significant determinants.

For awards, the explanatory power of the regressions is generally higher. In the regressions with the greatest explanatory power (R^2 greater than .8), NSF funding is the measure of research support, and additions, serials, and library staffing are all significant. Regardless of the research funding variable used, the library variables are significant in all but one regression.

Due to the small sample and missing data, these results are, again, indicative but not definitive. The implication, however, is that library resources affect research productivity even after research funding has been accounted for. This is an area that definitely requires further investigation.

The moderate correlations between research support (total and per faculty member) and the library resource variables are worth noting (Table 11).[5] One might have expected higher correlations, with a general pattern of high levels of

---

[4] Because 1989 publications per faculty member correlated best with 1987 library resources, Table 11 uses 1987 resource data.

[5] Table 11 reports 1989 library resources; correlations are similar regardless of the years used.

Table 11
Correlations between R&D and Science Library Resources

	1987 Science Library Resources			
	Additions	Serials	Holdings	Prof'l Staff
Total funding:				
NSF	-.02 (17)	.39 (13)	.23 (17)	.60 (18)
Federal	.11 (17)	.38 (13)	.31 (17)	.23 (18)
Total R&D	.27 (17)	.53 (13)	.49 (17)	.52 (18)
Funding per science faculty:				
NSF	-.25 (16)	.01 (12)	-.23 (16)	.11 (17)
Federal	-.10 (16)	-.04 (12)	-.11 (16)	-.27 (17)
Total R&D	-.04 (16)	.13 (12)	-.03 (16)	-.16 (17)

Table 12
Productivity Regressed on Library Resources and Research Funding

DEPENDENT VARIABLE	LIBRARY RESOURCES 1987				R&D $ PER SCI FACULTY			Adjusted R^2	N
	Addns	Holdings	Serials	Staff	NSF	Federal	Total		
Pubn's per science faculty member	x				***			.65	13
	x					*		.37	13
	x						***	.54	13
		x			***			.66	12
		*				**		.40	12
		x					**	.55	12
			x		**			.58	12
			*			x		.36	12@
			**				*	.60	12
				x	***			.63	16
				***		***		.56	16
				***			***	.63	16

All F statistics significant at <=.05 unless marked with @
x Entered in regression, not significant
*** Significant at <=.10**
**** Significant at <=.05**
***** Significant at <=.01**

Table 13
Productivity Regressed on Library Resources and Research Funding

DEPENDENT VARIABLE	LIBRARY RESOURCES 1987				R&D $ PER SCI FACULTY			Adjusted R^2	N
	Addns	Hold-ings	Serials	Staff	NSF	Federal	Total		
Awards per science faculty member	**				***			.86	13
	**					**		.59	13
	**						***	.62	13
		x			***			.83	12
		**				**		.52	12
		*					***	.60	12
			*		***			.82	12
			***			**		.71	12
			***				***	.78	12
				**	***			.85	16
				***		***		.70	16
				***			***	.66	16

All F statistics significant at <=.05 unless marked with @
x Entered in regression, not significant
*** Significant at <=.10**
**** Significant at <=.05**
***** Significant at <=.01**

institutional support for research activities and resources. Jones, Lindzey, and Coggeshall (1982), however, found moderate to low correlations between library size and university research expenditures. The available R&D figures, however, do not exactly match the library data: the NSF figures include social and behavioral sciences, and the other figures also include medicine, all of which were excluded from the library data. This finding requires further investigation.

DISCUSSION

The question addressed by this study is: is research productivity in science and engineering related to library resources?

Libraries may contribute to research directly or indirectly: directly by providing physical and intellectual access to the scholarly record, and indirectly by attracting faculty and graduate students who will produce research.

The library provides researchers with copies of publications and with information about publications and about the larger knowledge base. It helps researchers navigate the exploding knowledge base of their own and related fields. The library may play a particularly important role in supporting the work of graduate students, post-doctoral researchers, and other newcomers to the invisible college who are important partners in the university's research enterprise, some of whom will in time become major contributors in their own right.

The library's collection is an increasingly inadequate measure of its ability to access the published record. Databases, finding aids, resource sharing, information brokers, and document delivery services are among the many tools expanding the scope of research libraries and improving their services. We found data on these resources unavailable, however, forcing us to rely on more traditional collection measures.

The major finding of this study is that library resources in science and engineering are correlated with faculty productivity. The more science serials, the larger the science collections, and the more professional science librarians, the more publications and awards per faculty member.

Are these correlations simply a function of university characteristics that correlate with library resources? University size (measured by total faculty size) and research expenditures were tested. The results for this small sample are inconclusive, but it appears that after controlling for size and research expenditures, library resources still affect research productivity.

The library resource measures are themselves intercorrelated, making it difficult to sort out their relative influences. The greatest surprise may be the large correlation of research productivity with professional library staff (Tables 9, 12, and 13). Further investigation is needed to determine whether this is simply due to the correlation between staff and other resources. Many academic library users underestimate the role of professional staff in building and maintaining the collection, and in helping clients both to make maximum use of the collection and to locate information elsewhere.

This pilot study cannot demonstrate causality. Indeed; it indicates that empirical evidence of causal factors affecting university research may be hard to obtain. Science and engineering publication rates were quite stable over the five years of the study. If publication rates are relatively insensitive over the short term to environmental factors, this complicates research on possible determinants of publication rates, including but not limited to those connected to the library.

It is possible that the direction of causality is the reverse of our assumptions: that respected, productive institutions spend more on libraries because they have more to spend, or because of the prestige of having a major library. But these explanations seem unlikely, especially in times of stringent university budgets. The more logical direction of causality is for library resources to contribute to research productivity by tying researchers into the knowledge base. The two-year time lag in correlations between publications per capita and library resources tend to indicate that library resources do contribute to publications.

This study verifies librarians' claims of declining purchasing power. Given the correlations between library resources and publications, these declines could, in time, appear in publication rates.

IMPLICATIONS

The main implication of this study is that science library resources may well play a role in science research productivity. The possible effects of declining library purchasing power could be serious. Although this study cannot prove a causal connection, it does indicate a need for concern. By the time (if ever) a connection between science library resources and research is proven, it may be too late to undo the damage from missed acquisitions and deteriorating collections.

The implication for libraries is that while serials are important in the sciences (and many libraries have raided other parts of their budgets to maintain subscriptions in the face of price increases), so may be total holdings and especially professional staff. Unfortunately, without data on new resources like databases, we cannot speak to their value. The same is true of services like interlibrary loan, which enable even the smallest library to provide its users with virtually anything published. Unless a group like ARL defines the data elements and methods, such data won't be collected.

This was a pilot study, with all the shortcomings that the name implies. But the results indicate that a more extensive study, with a larger sample, would be justified for the following reasons:

- To sort out the effects of different library resources on research
- To investigate the effects of other variables on productivity,such as research expenditures
- To improve the data on science and engineering faculty and sort out the problem of including non-teaching researchers
- To collect better data on new types of information services and resources.

The authors of any future study should work closely with academic science librarians to develop measures for which data would be available. The best approach might be prospective, rather than retrospective, data collection to ensure comparable data.

Another area in need of investigation is the role of new information technologies, and especially the trade-offs between ownership of versus access to materials: is a library that owns little but acquires materials on request as useful for research as one with an extensive collection of its own?

Future studies at the level of discipline or subdiscipline are needed to develop production functions reflecting differences in the use of literature and of libraries across disciplines. Perceptual ratings of program quality or research contribution could then be included with objective measures.

This study provides some interesting and provocative findings, which should be used with caution given the following limitations:

- The results are valid only for the institutions in the sample, which comprises a small, select group of universities.
- The reliability of the library data are unknown.

- The short time frame and limited budget constrained the data collection and analysis.
- The disciplines covered by the dependent and independent variables did not coincide perfectly.

The results, however, demonstrate that the effect of science and engineering library resources on science and engineering research productivity is well worth further investigation.

REFERENCES

Akie, Ronald. "Periodical Prices 1986-88 Update." *Serials Librarian* 15, nos. 1/2 (1988): 43-59.

Association of Research Libraries. *Report of the ARL Serials Prices Project*. Washington, D.C.: ARL, 1989.

Association of Research Libraries. *ARL Statistics 1988-9*. Washington, D.C.: ARL, 1990.

Cartter, Allan M. *An Assessment of Quality in Graduate Education: a Comparative Study of Graduate Departments in 29 Academic Disciplines*. Washington, D.C.: American Council on Education, 1966.

Cohn, Steven F. "The Effects of Funding Changes upon the Rate of Knowledge Growth in Algebraic and Differential Topology, 1955-75." *Social Studies of Science* 16, no. 1 (1986): 23-60.

Cozzens, Susan E., ed. "Funding and Knowledge Growth." *Social Studies of Science* 16, no. 1 (1986b): issue.

Cozzens, Susan E. "Editor's Introduction." *Social Studies of Science* 16, no. 1 (1986a): 9-22.

Grannis, Chandler B. "Final Figures: Book Title Output and Prices." *Publishers Weekly* 234, no. 15 (1988): 68-72.

Grannis, Chandler B. "Titles and Prices, 1988: Final Figures." *Publishers Weekly* 236, no. 13 (1989): 24-27.

Holzner, Burkhart, William N. Dunn, and Muhammad Shahidullah. "An Accounting Scheme for Designing Science Impact Indicators." *Knowledge: Creation, Diffusion, Utilization* 9, no. 2 (1987): 173-204.

Jones, Lyle V., Gardner Lindzey, and Porter E. Coggeshall, eds. *An Assessment of Research-Doctorate Programs in the United States*. Washington, D.C.: National Academy Press, 1982.

Lenzini, Rebecca T. "Periodical Prices 1983-1985 Update." *Serials Librarian* 9, no. 4 (1985): 119-30.

"Library Materials Prices, 1989." *Library Issues* 10, no. 1 (1989).

Machlup, Fritz. *Information through the Printed Word: The Dissemination of Scholarly, Scientific, and Intellectual Knowledge*. 4 vols. New York: Praeger, 1978-80.

National Science Board. *Science and Engineering Indicators, 1989*. Washington, D.C.: Government Printing Office, 1989.

National Science Foundation. *Federal Support to Universities, Colleges, and Selected Nonprofit Institutions: Fiscal Year 1988*. Washington, D.C.: National Science Foundation, 1990.

Roose, Kenneth D., and Charles J. Andersen. *A Rating of Graduate Education Programs*. Washington, D.C.: American Council on Education, 1970.

Young, Peter. "Periodical Prices 1987-1989 Update." *Serials Librarian* 17, nos. 1/2 (1989): 11-37.

APPENDIX A

SURVEY FORM AND COVER LETTER

March 28, 1990

Dear :

Yours is one of a selected number of ARL Libraries we are asking to participate in an important study sponsored by the Council on Library Resources.

The purpose of the study is to explore the relationship between library expenditures and research productivity in science and engineering. Librarians know that the library plays a major role in research, yet little evidence exists to demonstrate this. Finding a link between library resources and research productivity will help to show the library's importance to research. We will combine the data that you give us with other data on on research productivity to test this relationship.

The Council on Library Resources is sponsoring this pilot study of a selected group of major research libraries to test the methods and hypotheses. With these results, the Council hopes to interest the National Science Foundation in more definitive research. Your cooperation in this study, therefore, may have a major payoff for libraries later on.

We want to make this as easy as possible for you. The data elements are from the ARL Library Statistics Questionnaire. We will get the totals for your library from the published ARL reports. We are asking you only for data on **science and engineering library resources and expenditures.** We realize that not every library collects these data, but many do, and others can develop them with a little effort. We hope that you agree that the importance of this study warrants the effort.

If you have the data in your own internal reporting format, if you prefer, just send us your reports (with Page 1 of the questionnaire) and we will convert it to our format.

Once again I want to stress the importance of your participation. Your library has been carefully selected to balance the sample's library and university characteristics. We need your help!

If you have questions, please feel free to call me or my research assistant, Diana Laslett, at 415-642-0855 (email to nav-lis at ucbcmsa.berkeley.edu; fax 415-642-5814).

Thank you for your help in this important study.

Yours truly,

Nancy Van House
Associate Professor

SCIENCE/ENGINEERING LIBRARY RESOURCES SURVEY

This survey collects data on library resources in science and engineering collections for 1984-5 through 1988-89.

Person filling this out:___________________________________**Phone:**____________________

Title: ____**Library:**______________________________

SCIENCE/ENGINEERING LIBRARY RESOURCES

Science and engineering are defined, for the purposes of this survey, as the **biological and physical sciences, mathematics, engineering, and computer science, not the social or behavioral sciences.**

Please do not include Medical Library statistics. If your Medical Library is nevertheless included in your science/engineering data, please check here___ and send us your data anyway.

If you cannot provide separate data on science/engineering library resources, check here___ and send back just this page (with name and address above). Thank you for your help.

Data elements are defined the same as for ARL Statistics whenever possible (definitions follow). To ease your burden, our line numbers correspond to ARL Questionnaire. (Some numbers missing because not all ARL items are included here.)

OR YOU CAN SEND...

If you can't answer all the questions, please answer as many as you can. We are most interested in expenditures. We would like 5 years of data, but if you have fewer, send what you have.

If you prefer, send us your library's own statistical summaries containing this information. We would rather you use our format, but we can live with yours. Check here___ and send us back this page with your summaries. **Be sure to tell us whom to call with questions.**

MAIL OR (PREFERABLY) FAX TO

Nancy Van House
School of Library and Information Studies
University of California
Berkeley, CA 94720
phone 415-642-0855
fax 415-642-5814

QUESTIONS? Call Nancy Van House or Diana Laslett at the number above.

Please return by
April 13, 1990

Thanks very much for your help!

Please return this page with the next two.

SCIENCE/ENGINEERING LIBRARY RESOURCES SURVEY

Please enter your library's data FOR SCIENCE AND ENGINEERING for each year.

****We are most interested in materials expenditures.**

Line numbers and definitions are the same as ARL Statistics Questionnaire EXCEPT those with letters (e.g., X, XX, etc.)

DATA ELEMENT Science/Engineering ONLY!	1984-5	1985-6	1986-7	1987-8	1988-9
COLLECTIONS 4.Volumes added during year -- net (exclude microforms, uncataloged govt. docs, maps, a/v materials)					
5.Volumes held, as of June 30.					
W. Other materials added, including technical reports, government documents, microforms, not included on line 4					
WW. Other materials held, June 30.					
9.Total number of current serials received (purchased and not purchased), incl. exchanges, gifts, etc.					
PERSONNEL 11.No. of professional staff, FTE, whose primary responsibility is science/engineering, including public services and collection development					

SCIENCE/ENGINEERING LIBRARY RESOURCES SURVEY

Please enter your library's data FOR SCIENCE AND ENGINEERING for each year.

****We are most interested in materials expenditures.**

Line numbers and definitions are the same as ARL Statistics Questionnaire EXCEPT those with letters (e.g., X, XX, etc.)

DATA ELEMENT Science/Engineering ONLY!	1984-5	1985-6	1986-7	1987-8	1988-9
****EXPENDITURES** **15.Monographs (expenditures for volumes reported on line 4)**					
16.Current serials including periodicals (exp. for purchases from line 9)					
X.Exp. from library budget for doc. delivery services, incl. info brokers, ISI, NTIS, etc., not incl. on lines 15 or 16					
19.Total science/engineering library materials exp.(monographs, serials, other library materials, misc., not incl. binding)					
Y.Total expenditures on online search services. Include subscription and connect charges for vendor services and CD-ROM products. Do not include equipment.					
MISC. Z.No. of science/engineering interlibrary borrowing requests, plus items acqd. from info brokers, ISI, etc.					
ZZ.Science/engineering faculty, FTE					

IV B-41

FOR REFERENCE----USE THESE DEFINITIONS WHEN COMPLETING QUESTIONNAIRE

ARL STATISTICS, 1988-89

Instructions for Completing the Questionnaire

Definitions of the statistical categories used in this questionnaire can be found in *American National Standard for Library and Information Sciences and Related Publishing Practices - Library Statistics*. Z39.7-1983. (New York, American National Standards Institute, 1983.)

The questionnaire assumes a fiscal year ending June 30, 1989. If your fiscal year is different, please provide a footnote on p. 4 of the questionnaire.

Please do not use decimals. All figures should be rounded to the nearest whole number.

Please do not leave any lines blank. If the appropriate answer is zero or none, use **0**. If an exact figure is unavailable, use **U/A**. If a question is not applicable to your library, use **N/A**.

In a university that includes both main and branch campuses, an effort should be made to report figures for the main campus only. (The U.S. National Center for Education Statistics in its Integrated Postsecondary Education Data System (IPEDS) survey described a **branch campus** as one "located in a community different from that of its parent institution ... beyond a reasonable commuting distance from the main campus ... The educational activities at the location must be organized on a relatively permanent basis ... and include course offerings for one or more complete college-level programs of at least one full year.") If figures for libraries located on branch campuses are reported, please provide an explanatory footnote.

A **branch library** is defined as an auxiliary library service outlet with quarters separate from the central library of a system, which has a basic collection of books and other materials, a regular staffing level, and an established schedule. A branch library is administered either by the central library or (as in the case of some law and medical libraries) through the administrative structure of other units within the university. Departmental study/reading rooms are not included. Do include branch data.

Questions 4-5 Collections. Use the ANSI Z39.7-1983 definition for **volume** as follows:

> a physical unit of any printed, typewritten, handwritten, mimeographed, or processed work, contained in one binding or portfolio, hardbound or paperbound, which has been cataloged, classified, and made ready for use.

Include duplicates and bound volumes of periodicals. Exclude microforms, maps, nonprint materials, and uncataloged items. Exclude government document volumes unless they are cataloged, classified, and shelved as part of the general collection. (Data on documents holdings are requested in the Supplementary Statistics questionnaire.) If any of these items cannot be excluded, please provide an explanatory footnote on the questionnaire. For purposes of this questionnaire, unclassified bound serials that are arranged in alphabetical order are considered classified.

Question 4 Report net number of volumes purchased. Include all volumes for which an expenditure was made during 198 -8 , including volumes paid for in advance but not received during the fiscal year. Include monographs in series and continuations. If only number of titles purchased can be reported, please report the data and provide an explanatory footnote on the questionnaire. **Note:** This question is concerned with volumes purchased rather than volumes received or cataloged. Question 15 requests the expenditure for the volumes counted here.

Questions 9. Serials. Include duplicate subscriptions and only those government document serials that are cataloged, classified, and shelved as part of the general collection; do not include government documents serials that are housed in a separate government documents collection. (Data on government documents serials is requested in the Supplemental Statistics questionnaire.) Exclude monographic and publishers' series. A **serial** is

> a publication issued in successive parts, usually at regular intervals, and as a rule, intended to be continued indefinitely. Serials include periodicals, newspapers, annuals (reports, yearbooks, etc.), memoirs, proceedings, and transactions of societies.

Questions 11 Personnel. Report the number of staff in filled positions, or positions which are only temporarily vacant. Include cost recovery positions and staff hired for special projects and grants, but provide an explanatory footnote indicating the number of such staff. If such staff cannot be included, provide a footnote on p. 4 of the questionnaire. The number of FTE staff should be determined on the basis of the length of the work week in the reporting library. **Round figures to the nearest whole numbers.**

Question 11. Since the criteria for determining professional status vary among libraries, there is no attempt to define the term "professional." Each library should report those staff members it considers professional, including, when appropriate, staff who are not librarians in the strict sense of the term, such as computer experts, systems analysts, budget officers, etc.

Questions 15- 19 Expenditures. Report all expenditures of funds that come to the library from the regular institutional budget and from other sources, such as research grants, special projects, gifts and endowments, and fees for service. Canadian libraries should report expenditures in both Canadian and U.S. dollars. To determine figures in U.S. dollars, divide Canadian dollar amounts by 1.2026, the average monthly noon exchange rate published in the Bank of Canada *Review* for the period July 1988-June 1989. **Please round figures to the nearest dollar.**

Question 15. Include here expenditures for volumes reported on line 6.

Question 16. Exclude monographic and publishers' series.

Question Y. Include net fees paid from the library budget to search service vendors for connect time, subscriptions, and CD-ROM products. Do not include costs passed through to users, or equipment.

Question Z. Interlibrary borrowing requests for science/engineering materials or users, plus items acquired from other sources such as information brokers, ISI, NTIS, etc., for your users.

APPENDIX B
Materials Price Indexing

Library materials prices are going up much faster than the cost of living (Library Materials Prices, 1989). A common deflator like the Consumer Price Index would understate the decline in library purchasing power. A library materials price index is needed instead.

Price increases vary according to type of material, subject, and country of publication. It is difficult, therefore, to come up with the ideal price index for all the libraries in our sample, since the mix of materials purchased will differ. The following, however, corrects for average increases in the prices of serials and monographs.

Serials Prices

For serials prices, this study used the annual studies done by The Faxon Company from their extensive database (Lenzini, 1985; Akie, 1988; and Young, 1989). The figures used were the average prepaid annual subscription price per title for serials in the Science Citation Index, weighted according to the relative number of Faxon subscriptions.[6] The SCI titles represent the closest category in the Faxon studies to science and engineering periodicals.

Year	Average Price	% change from year before	% change since 1985
1985	$180.51	--	--
1986	207.48	14.9%	14.9%
1987	251.68	21.3	39.4
1988	274.74	9.2	52.2
1989	299.96	9.2	66.2

Using these indexes, average serials expenditures among the libraries in our sample declined in 1985-6 and 1987-8, with 1988-9 remaining below the 1984-5 level (Table 4).

---

[6] Faxon is a major periodicals subscription agent for U.S. libraries. The weights correct for the extent to which journals are actually purchased by U.S. libraries.

Monograph Prices

Monograph price indexes were taken from the annual Publishers Weekly analysis of American book prices (Grannis, 1989; Grannis, 1988). The figures used were the average price per title for hardcover domestic and imported science books. Without knowing the proportion of foreign monographs among our samples' acquisitions, any correction for foreign acquisitions is somewhat arbitrary; PW's index includes all the domestic and imported titles listed in PW's weekly record listings.

The 1989 figure will not be available until approximately October of 1990. The figure used here was extrapolated by assuming the same percentage increase from 1988 to 1989 as from 1987 to 1988.

Again, we find real expenditures declining, then increasing, during the time period examined, with real expenditures at the end of five years below where they began (Table 5).

Year	Average Price	% change	% change since 1985
1985	$51.19	--	--
1986	55.65	8.7%	8.7%
1987	62.16	11.7	21.4
1988	66.91	7.6	30.7
1989	72.00 (est.)	7.6	40.7

APPENDIX C
Correlations of Science Library Resources with Productivity Measures by Year

1985

	PYIs 1985	Publications 1985
Net Additions, 1985	.33 (14)	.21 (14)
Holdings, 1985	.37 (16)	.40 (16)
Serial Subscriptions, 1985	.64 (13)	.49 (13)
Prof'l Staff FTE, 1985	.70 (20)	.84 (20)
Expenditures, 1985 $		
Monographs	.46 (21)	.64 (21)
Serials	.70 (20)	.64 (20)

1986

	PYIs 1986	Publications 1986
Net Additions		
1985	.10 (14)	.22 (14)
1986	.07 (18)	.19 (18)
Holdings		
1985	.10 (16)	.41 (16)
1986	.07 (17)	.29 (17)

Appendix C (page 2)
Correlations of Science Library Resources with Productivity Measures by Year

	1986 (cont)	
	PYIs 1986	Publications 1986
Serial Subscriptions		
1985	.22 (13)	.50 (13)
1986	.18 (14)	.47 (14)
Prof'l Staff FTE		
1985	.35 (20)	.85 (20)
1986	.34 (21)	.80 (21)
Expenditures, 1985 $		
Monographs		
1985	.49 (21)	.64 (20)
1986	.42 (23)	.48 (23)
Serials		
1985	.37 (20)	.63 (20)
1986	.42 (21)	.62 (21)

Appendix C (page 3)
Correlations of Science Library Resources with Productivity Measures by Year

	1987	
	PYIs 1987	Publications 1987
Net Additions		
1985	.34 (14)	.21 (14)
1986	.24 (18)	.19 (18)
1987	.07 (18)	.08 (18)
Holdings		
1985	.35 (16)	.39 (16)
1986	.36 (16)	.28 (16)
1987	.28 (18)	.31 (18)
Serial Subscriptions		
1985	.55 (13)	.50 (13)
1986	.52 (14)	.47 (14)
1987	.54 (14)	.48 (14)
Prof'l Staff FTE		
1985	.68 (20)	.85 (20)
1986	.62 (21)	.80 (21)

Appendix C (page 4)
Correlations of Science Library Resources with Productivity Measures by Year

1987 (cont)

	PYIs 1987	Publications 1987
Prof'l Staff FTE		
1987	.63 (21)	.82 (21)
Expenditures, 1985 $		
Monographs		
1985	.37 (21)	.62 (21)
1986	.26 (23)	.47 (23)
1987	.35 (23)	.41 (23)
Serials		
1985	.63 (20)	.64 (20)
1986	.55 (21)	.61 (21)
1987	.57 (22)	.58 (22)

Appendix C (page 5)
Correlations of Science Library Resources with Productivity Measures by Year

1988

	PYIs 1988	Publications 1988
Net Additions		
1985	.37 (14)	.22 (14)
1986	.41 (18)	.20 (18)
1987	.31 (18)	.10 (18)
1988	.17 (18)	.03 (18)
Holdings		
1985	.48 (16)	.40 (16)
1986	.38 (17)	.30 (17)
1987	.39 (18)	.32 (18)
1988	.33 (18)	.26 (18)
Serial Subscriptions		
1985	.73 (13)	.51 (13)
1986	.65 (14)	.49 (14)
1987	.66 (14)	.49 (14)
1988	.68 (14)	.54 (14)

Appendix C (page 6)
Correlations of Science Library Resources with Productivity Measures by Year
1988 (cont)

	PYIs 1988	Publications 1988
Prof'l staff FTE		
1985	.63 (20)	.84 (20)
1986	.57 (21)	.79 (21)
1987	.56 (21)	.81 (21)
1988	.42 (23)	.41 (23)
Expenditures, 1985 $		
Monographs		
1985	.60 (21)	.64 (21)
1986	.52 (23)	.48 (23)
1987	.54 (23)	.42 (23)
1988	.40 (24)	.41 (24)
Serials		
1985	.67 (20)	.65 (20)
1986	.60 (21)	.62 (21)
1987	.61 (22)	.59 (22)
1988	.65 (23)	.64 (23)

Appendix C (page 7)
Correlations of Science Library Resources with Productivity Measures
by Year
1989

	NAS Members	NAE Members	PYIs 1989	AWARDS* 1985-9	Pubs 1989
Net Additions					
1985	.21 (14)	.34 (14)	.50 (14)	.31 (14)	.21 (14)
1986	.27 (18)	.32 (18)	.31 (18)	.31 (18)	.20 (18)
1987	.09 (18)	.14 (18)	.23 (18)	.13 (18)	.09 (18)
1988	-.07 (18)	.02 (18)	.20 (18)	-.04 (18)	.02 (18)
1989	.10 (20)	.19 (20)	.25 (20)	.16 (20)	.34 (20)
Holdings					
1985	.39 (16)	.44 (16)	.58 (16)	.43 (16)	.41 (16)
1986	.25 (17)	.37 (17)	.46 (17)	.34 (17)	.31 (17)
1987	.27 (18)	.36 (18)	.41 (18)	.33 (18)	.33 (18)
1988	.21 (18)	.31 (18)	.38 (18)	.28 (18)	.27 (18)
1989	.21 (22)	.29 (22)	.41 (22)	.28 (22)	.40 (22)

*NAS + NAE + PYIs 1985 through 1989

Appendix C (page 8)
Correlations of Science Library Resources with Productivity Measures by Year

1989 (cont)

	NAS Members	NAE Members	PYIs 1989	AWARDS 1985-9	Pubs 1989
Serial Subscriptions					
1985	.54 (13)	.44 (13)	.71 (13)	.54 (13)	.49 (13)
1986	.47 (14)	.41 (14)	.62 (14)	.49 (14)	.47 (14)
1987	.48 (14)	.42 (14)	.64 (14)	.51 (14)	.48 (14)
1988	.55 (14)	.47 (14)	.67 (14)	.57 (14)	.53 (14)
1989	.44 (22)	.41 (22)	.64 (22)	.47 (22)	.46 (22)
Prof'l Staff FTE					
1985	.69 (20)	.65 (20)	.67 (20)	.71 (20)	.84 (20)
1986	.61 (21)	.62 (21)	.55 (21)	.64 (21)	.80 (21)
1987	.61 (21)	.62 (21)	.57 (21)	.64 (21)	.82 (21)
1988	.24 (23)	.24 (23)	.47 (23)	.29 (23)	.41 (23)
1989	.20 (26)	.22 (26)	.44 (26)	.25 (26)	.35 (26)

Appendix C (page 9)
Correlations of Science Library Resources with Productivity Measures by Year

1989 (cont)

	NAS Members	NAE Members	PYIs 1989	AWARDS 1985-9	Pubs 1989
Expenditures, 1985 $					
Monographs					
1985	.57 (21)	.48 (21)	.58 (21)	.55 (21)	.65 (21)
1986	.47 (23)	.42 (23)	.37 (23)	.46 (23)	.49 (23)
1987	.46 (23)	.47 (23)	.35 (23)	.48 (23)	.44 (23)
1988	.38 (24)	.37 (24)	.18 (24)	.37 (24)	.42 (24)
1989	.28 (27)	.27 (27)	.23 (27)	.29 (27)	.40 (27)
Serials					
1985	.71 (20)	.64 (20)	.70 (20)	.71 (20)	.65 (20)
1986	.65 (21)	.56 (21)	.54 (21)	.63 (21)	.62 (21)
1987	.64 (22)	.55 (22)	.57 (22)	.63 (22)	.59 (22)
1988	.66 (23)	.55 (23)	.58 (23)	.65 (23)	.64 (23)
1989	.62 (25)	.58 (25)	.57 (25)	.64 (25)	.58 (25)